DOCUMENT RESUME

ED 267 396 CS 209 264

AUTHOR Hayes, John R.; And Others
TITLE Cognitive Processes in Revision. CDC Technical Report No. 12.
INSTITUTION Carnegie-Mellon Univ., Pittsburgh, PA. Communications Design Center.
PUB DATE Apr 85
NOTE 89p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Behavior Patterns; *Cognitive Processes; *Learning Strategies; Learning Theories; *Models; *Protocol Analysis; Reading Comprehension; *Revision (Written Composition); Rhetorical Invention; Writing Evaluation; *Writing Processes
IDENTIFIERS *Writing Strategies

ABSTRACT

A new model of the revision process in written composition, based on the results of thinking aloud protocol studies, is presented in this report. The report begins by discussing earlier observations and theories of revision that establish four points: (1) there are large differences among writers in the amount of revising they do, (2) expert revisors attend to more global revising problems than do novices, (3) writers have more difficulty detecting faulty referring expressions when revising their own text than when revising the texts of other writers, and (4) the ability to detect text problems appears to be separate from the ability to fix these problems. The report then describes the model according to the following central subprocesses: task definition, evaluation for problem detection and diagnosis, and strategy selection. The report concludes with a detailed discussion of the protocol data on which the model is based. References, tables of data, a sample revision task, and an explanation of talk-aloud protocols of the revision process are appended. (HOD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED267396

CDC Technical Report No. 12

# Cognitive Processes in Revision

John R. Hayes, Linda Flower, Karen Schriver, James Stratman, and Linda Carey

April 1985

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Thomas M. Duffy

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Communications Design Center

Carnegie-Mellon University
160 Baker Hall
Pittsburgh, PA 15213

CS209264

The Communications Design Center was established in 1979 to solve communication problems in business, industry, government, and the professions.

The goals of the Center are:

- to encourage interdisciplinary approaches to the solution of visual and verbal communication problems;
- to promote cooperation between communication researchers at the University and those outside the University who have practical interests in communication problems;
- to provide resources and an organizational setting for research in visual and verbal communication;
- to train professional writers, researchers, and teachers;
- to promote the development of educational materials for communication programs on and off the CMU campus.

This report is one of a series from our on-going research on communication and document design. If you have questions or comments on this report, or if you would like to find out more about the activities of the CDC, contact:

Director
Communications Design Center
Baker Hall 160
Carnegie-Mellon University
Pittsburgh, PA 15213
(412) 578-2906

# Cognitive Processes in Revision

**Technical Report to**

**the National Science Foundation**

John R. Hayes

Linda Flower

Karen A. Schriver

James Stratman

Linda Carey

Carnegie-Mellon University

October 1984

To appear in Advances in Applied Psycholinguistics, Volume II: Reading, Writing, and Language Processing, edited by Sheldon Rosenberg, Cambridge University Press, Cambridge, England, in press.

## SUMMARY

In this paper we present a new conceptual integration and elaboration of the cognitive processes involved in revision in a process model. Our major findings, organized around the central subprocesses are the following:

*Task Definition* — Due to the ill-defined nature of writing tasks, differences in *task definition* contribute importantly to the differences in performance between expert and novice revisors. One important aspect of the difference in definition is that experts see revision as a whole text task while students see it largely as a sentence level task. Experts perform better than novices not just because they perform certain subskills better than novices, but also and more importantly because they are doing a better task--that is, one better suited to improving text.

*Evaluation* — One part of the revision model is a new model of *evaluation*--one that is different than dissonance models, applying to the evaluation of *plans* as well as text. Since revision can be carried on in the absence of text, it can not depend entirely on the comparison of text and intention. Our model differs in this respect from the models of Sommers, of Bridwell, and of Scardamalia and Bereiter.

*Problem Representation* — *Problem representation* constitutes a continuum from *ill-defined to well-defined*. *Detections* which fall on the ill-defined end contain little information; *diagnoses*, on the other end, are based on concept recognition--they categorize the problem and gives one access to appropriate procedures.

*Detection* — Novices persistently fail to perceive text problems that experts detect easily. Experts consistently outperform novices both in the detection of local and global text problems.

*Diagnosis* — *Diagnoses* vary in quality. The diagnoses of experts are more global, elaborated, integrated and procedural than those of novices.

*Strategy Selection* — *Strategy selection* in revision involves both strategies for managing the revising process: (*ignore, delay* and *search*) and for modifying text: (*rewrite* or *revise*). Experts tend to exercise the choice among these options more effectively than do novices. The *means-ends repertory* of expert revisors is more elaborated and finely partitioned than that of novices. Most strikingly, it includes more goals concerning global text problems, and it contains *means* absent from novice repertories for solving these problems.

# I. INTRODUCTION

The primary purpose of this report is to present a new model of the revision process in written composition--a model based on the results of thinking aloud protocol studies. We will proceed first by discussing earlier observations and theories of revision. Then we will describe the model and, finally, we will discuss in detail the protocol data on which the model is based.

## Earlier Observations

While there have been relatively few empirical studies on revision, those that are available provide valuable guideposts for the development of theory. The literature reviewed below establishes four points:

1. There are large differences among writers in the amount of revising they do. Expert's revise more than novices.
2. Expert revisors attend to more global revising problems than do novices.
3. Writers have more difficulty detecting faulty referring expressions when revising their own text than when revising the texts of other writers.
4. The ability to detect text problems appears to be separate from the ability to fix these problems.

Writers differ greatly in the amount of revising they do.

Murray (1978) reviewing testimony of eminent writers, concludes that for them, "writing is rewriting." Here, Murray is describing the propensity of skilled writers to spend far more time in revision than in the production of the original draft. On the other hand, Bracewell, Scardamalia, and Bereiter (1978) found that fourth graders hardly revise at all, that 8th graders' revisions hurt more than they help, and that for 12th graders, helpful revisions narrowly outnumbered harmful ones. Bridwell's results were a bit more positive about 12th graders revisions than were those of Bracewell et al's (1978). She found that 12th graders second drafts were considerably better in "general merit" and mechanics than their first drafts. Pianko (1977) reported that college freshmen devote less than 9% of their composing time to rereading and revising. Clearly, writers differ widely in the amount they revise. In general, it appears that the more expert the writer, the greater proportion of writing time the writer will spend in revision.

Experts attend more to global problems than do novices.

The literature suggests that experts and novices define revising in very different ways. Broadly, revision may be defined as the writer's attempt to improve a plan or text. Within this definition, experts appear to attend systematically to different aspect of the text than do novices.

Stallard (1974) found that only 2.5% of 12th graders revisions were focused above the word and sentence level. Bridwell (1980) who also studied 12th graders, found about 11% of revisions above the sentence level. Bridwell's figure of 11% may be higher than Stallard's because Bridwell counted each sentence in a multi-sentence revision as a separate occurrence of a multi-sentence revision. If Bridwell had counted each multi-sentence revision just once for each group of revised sentences as Stallard did, the percentage of such revisions could not exceed 5.9%, (if each multi-sentence revision consisted of just two sentences) and would be lower if some revision involved 3 or more sentences.

Beach (1976), studying college juniors and seniors, found that students who revised extensively "tended to conceive of the paper in holistic terms" and to infer "general patterns of development." Students who did not revise extensively "evaluated only separate bits" of their papers.

Sommers (1978) found that college freshmen "understand the revision process as a rewording activity . . . .They concentrate on particular words apart from their role in the text." In contrast, experienced writers, e.g., journalists, editors, and academics "describe their primary objectives when revising as finding the form or shape of their argument." Further, Sommers (1978) found that "the experienced writers have a second objective; a concern for their readership." Moreover, Sommers says, "At the heart of revision is the process by which writers recognize and resolve the dissonance they sense in their writing."

Faigley and Witte (1983), who studied changes in meaning resulting from revision, found that experts were more likely to change meaning through revision than were novices. They observed that the revisions of inexperienced college writers, resulted in changed meaning in 12% of cases; the revisions of experienced college writers, in 25% of cases, and the revisions of expert adult writers, in 34% of cases.

These results suggest that expert revisions define the task of revision as being more global and more focused on meaning and audience than do novices.

Writers have difficulty detecting faults in their own text.

Bartlett (1981) compared revision processes in fifth grade students who were revising both

their own and other writers' texts. In one study, she found that when 5th grade children were revising their own texts, they were able to find 56% of missing subjects or predicates, but only 10% of faulty referring expressions. In contrast, when the children were revising the text of other writers, they detected about half of each type of problem. Knowledge of their own intention as writers apparently made it difficult for them to detect faulty references in their own texts.

Finding and fixing problems appear to be separate skills.

In a second study, Bartlett asked 6th and 7th graders to revise paragraphs which embodied one of two types of reference problems. The first type of problem involved sentences of the form "A boy named Joe. . . Another boy named Bill. . . He . . . ." The second involved sentences of the form "A boy. . . Another boy. . . He. . . ."

Confronted with the first problem type, the children were able to detect 62% of the problems and then were able to correct 95% of those they detected. With the second problem type, they detected 52% of the problems (about the same as before) but were able to correct only 58% of those they detected. Solution to the second type of problem, then, appeared to be limited in two ways. First, subjects were able to detect only half of the problems and, of course, they did not solve those problems they did not detect. Thus, solution was limited by the subjects' ability to *detect* problems. Second, subjects were able to solve only about half of the problems that they did detect. Thus, solution was also limited by the subjects ability to solve identified problems. Solution of the first type of problem appeared to be limited just by the ability to detect problems since subjects solved almost all of the problems they detected. However, we have no evidence that the subjects could have solved the problems that they failed to detect if they had been pointed out to them. Thus, solution of the first problem type was certainly limited by the subjects' ability to detect problems but may also be limited by their ability to solve them.

Bereiter and Scardamalia (1983) have suggested that 6-8th graders ability to revise is limited much more by their ability to fix problems than by their ability to detect them. On the basis of presently available data, our best guess is that both the ability to detect problems and the ability to fix them once they are detected act independently to limit the students ability to revise.

## Earlier Theory: Cues Which Initiate Revision

Several researchers have suggested that the cue or initiating condition for revision is a dissonance or incongruity between intention and execution. For example, Bridwell (1980) suggests then when rereading the text, "the writer may either verify what is on the page or perceive some dissonance." Perception of dissonance is the cue which may lead to a decision to change the

text. According to Sommers (1978), "the anticipation of a reader's judgment causes a feeling of dissonance when the writer recognizes incongruities between intention and execution." It is this recognition according to Sommers that leads the writer to make revisions.

The most complete model of revision to date is Scardamalia and Bereiter's (1983) C.D.O. model (for Compare, Diagnose, and Operate). According to Scardamalia and Bereiter, "During the course of composition, two kinds of mental representations are built up and stored in long-term memory. These are a representation of the text as written up to the time, and a representation of the text as intended. The C.D.O. process is initiated by a perceived mismatch between these two representations."

While the idea that revision is initiated by the discovery of dissonance between intention and execution is attractive, we believe that it is not adequate to account for all of the phenomena we observe in revision. To make it clear why we think so, we will first define some necessary concepts. When we speak of intention, we refer to the author's writing plan to produce a text which will accomplish a purpose such as conveying facts or convincing an audience. This writing plan or network of working goals was itself constructed out of the writer's knowledge of goals, plans, constraints, and criteria for discourse and for problem solving in general. By text, we mean the external written product which the writer produces (with greater or lesser skill) in an attempt to carry out the writing plan. If the text is reasonably well written, a reader can usually infer the major aspects of the writing plan from the text alone. The writing plan, then, while it is most available to the writer, is also available in various degrees to others.

With these definitions in mind, we see that the revisor may engage in evaluations of several kinds:

1. They may evaluate the text against general criteria for texts such as standards of spelling, grammar, and clarity. Such evaluations can be carried out even if the revisor is not aware of the writer's plans for this particular text. Certainly, a text will fail these general criteria if it is so badly garbled that the reader is completely unable to decipher the writer's intention.

2. They may evaluate the text against the writer's intention. When writers revise their own texts, they frequently find that the text doesn't fully express their intended meaning. They may find that the words and examples that they have chosen, while they are close to the intended meaning, are not fully accurate and need to be revised. When the revisors are not the writer, they must infer the writers' plan from the text. The inferred plan can then be compared to the text. For example, the revisor may say, "I see what he is trying to do, but he needs a better example."

3. They may evaluate the plan against criteria which they believe plans should meet. For example, a writer might plan to present a list of examples supporting a point and then, reflecting on the plan, decide that the listing approach wouldn't be convincing because it is boring. Another interesting case is one in which the writer plans an argument and then executes it in writing. On reading the text, he finds that it

accurately represents the plan but that the plan is built on contradictory assumptions. Thus, the text matches the plan (a positive type 2 evaluation) but the plan itself is rejected.

Of these three types of evaluation, the second involves the comparison of text and intention but the first and third do not. Of the three, we believe that the third is the most important in producing high quality revisions. In fact, we propose that the greater power of expert revisors comes not from better comparison of text and intention, but rather from better evaluation of plans with respect to certain general goals writers have for plans.

## II. THE MODEL

In the present section we introduce a theoretical model of the revision processes we observed. However, before we explicate the specific features of this model, we feel it is important to address several questions about the value and place of such models in research and instruction. In particular, we wish to explain how models induced from protocol data help us to create more testable theories of writing task performance. Recently, model-building based upon protocol data has been attacked as a poor method of understanding underlying cognitive skills in writing (e.g., Cooper & Holzman, 1983, 1985). Indeed, because the recent essay by Cooper and Holzman (1985) contains most of the serious (and recurrent) criticisms leveled at protocol methodology, we will here briefly address the objections they raise.

### Goals And Criticisms Of Protocol-Derived Models

Three criticisms of the use of protocols to study writing have been advanced by Cooper and Holzman. We will identify each one here, then respond to each in detail below.

1. Valid empirical research can only begin with a "well-defined" object of research. Cognitive processes in writing are ill-defined; therefore, using protocols to study these processes is impossible until the precise nature of these processes is clearly, and unambiguously explicated.

2. Writing theories induced from protocol data are not testable in principle, because writing processes are too complex to capture in protocols and because the protocol verbalizations themselves do not provide any information about these writing processes.

3. Experts and novices differ in their capacity to give reports of their activities while writing; therefore, comparisons of experts and novices are likely to reflect differences in this capacity, not necessarily differences in particular skills or processes.

Let us take up each of these criticisms in turn.

(1) Cooper and Holzman and other observers assume that the purpose of protocol methodology in our studies is to test - or validate - a theory of cognitive processes in writing. In addition, these observers have contended that we have not sufficiently defined our "object of research" for purposes of such testing. For instance, Cooper and Holzman argue that before using the methodology, we should be able to define the difference between the cognitive processes underlying writing and those underlying speech generation in an oratorical context. However, two unfortunate (and, we feel, mistaken) assumptions lie at the base of such criticisms.

First, so far in our research we have used protocol methodology primarily to construct and enrich theories of writing task performance, not to test our theories or models. The problem with current research on writing as we perceive it is that, where underlying cognitive processes are concerned, there has been no strong, or "well-defined" theory to test - certainly none before Emig (1970). To that extent, we agree with the concern of Cooper and Holzman for creating testable theories. But because we have noted the absence of empirically-based process theories, we began building theories using the untried observational method of protocols. The goal of such theory-building, in contrast with theory-testing (or hypothesis-testing), is to locate processes that might be more carefully studied, that is, to locate and explore phenomena not previously observed that might have theoretical significance. The research available indicates that protocols are ideal for this theory-building purpose.

However, observers like Cooper and Holzman (1985) appear to demand that phenomena must be perfectly understood before they can be either observed or tested. They state, "As a theory is useful to science only when applied to a well-defined object of research, so a methodology is useful only to the extent that it can be used to test a carefully articulated theory" (p. 98). We reject both assertions. They are based upon a conception of the relation between methodology and theory that is unnecessarily narrow - one that would strangle the way many scientists do productive research. In the first instance, many a theory has been useful to science when the object of that theory was only partially defined - Freud's theory of dreams is just one powerful case in point. (Whatever else his current detractors will say, we doubt if they will say his theory is "useless.") In the second instance, many an important theory has been built (or discovered) using a different methodology than the one used subsequently to test or validate its hypotheses. To make the kind of sweeping claim that Cooper and Holzman do (above) is to blithely ignore the fact that there may be separate methods each for building, testing and interrelating theories.

Indeed, our situation with protocol methodology is somewhat analogous to the scientific situation that emerged with the invention of microscopes. As is commonly known, early microscopes were initially important as means of generating new theories of disease etiology - theories that no one quite knew how to test or validate at the time, but which stimulted an

entirely new direction of hypothesis-testing. In a similar way, we are using protocols to make novel but (as far as possible) systematic observations about the information-processing activities of writers. Our search for interpretations of what we observe necessarily proceeds in parallel with the search for a better (i.e., more testable) model or theory (see Simon & Ericsson (1985), pp. 5 - 10.).

(2) A second pair of criticisms is also frequently voiced. One is that the writing theories induced from protocol data are not testable in principle (Cooper & Holzman, 1985, p. 98). For instance, claiming Simon and Ericsson as support, Cooper and Holzman state that "laboratory protocol methodologies cannot deal with complicated cognitive processes," and that "common experience shows that cognitive processes in non-laboratory conditions are also unreachable in this fashion" (p.99). Both statements are wholly unfounded (Cooper and Holzman offer not a single study in support), and reflect a superficial reading (at best) of the research on protocol methodology reported by Simon and Ericsson. Indeed, Cooper and Holzman's statements disregard the thrust of Simon and Ericsson's entire book. For while Simon and Ericsson go to great lengths to describe limitations of protocol methodology for specific purposes and under specific conditions, they never state prima facie that it cannot deal with complicated cognitive processes. On the contrary, they conclude, after a massive review of empirical studies, that such a criticism is unfounded (see especially Chapters 4 and 5). Protocol data is shown to be a powerful means of both testing and building theories - not merely theories about simple processes, but also theories about higher-level processes as well. As they plainly state, "higher-level activities like reading and writing are made up of numerous sub-processes that use short-term memory for inputs and outputs. (Protocol) subjects' awareness of these inputs and outputs enables them to report the occurrence and duration of higher-level activities" (p. 243).

As for tracing cognitive processes outside of the laboratory, numerous successful protocol studies of subjects executing "real" complex tasks have been conducted (see the bibliography in Simon and Ericsson, 1984). The evidence produced by these studies indicates that many supposedly "unreachable" or "unreportable" cognitive processes can be either reported or inferred from transcripts. Many "ecologically valid" studies are also discussed by Simon and Ericsson.

Another criticism often closely follows the preceding one. This criticism is that the verbal data produced by protocols has nothing to do with the cognitive processes in question (i.e., with cognitive processes in writing). In other words, the protocol data is merely "epiphenomenal." For example, Cooper and Holzman (1985) contend that protocol methodology merely "produces interesting narratives, data for a theory about what certain writers will say about the writing process, but not data about the writing process, nor even, given its limitations, particularly useful data about the cognitive processes of writers in (a) particular situation" (p. 99.) But this charge that verbal reports of writers are epiphenomenal has little or no empirical basis. They offer

(again) not one study in support of their claims. Indeed, this criticism is also carefully addressed by Simon and Ericsson, and it would appear Cooper and Holzman ignored their discussion or failed to comprehend it. As Simon and Ericsson note, "Whenever verbalizations correspond to plausible intermediate states in a processing model for the problem solving activity, we can plausibly infer that this information is actually used in generating the problem solution" (p. 171, our underline). In addition, Simon and Ericsson describe three criterion for assessing the relevance of verbalizations to a given task.

- The vebalizations should be relevant to the given task.
- The verbalizations, to be pertinent, should be logically consistent with the vebalizations that just precede them.
- A subset of the information heeded during task performance will be remembered.

We take our protocols to satisfy all three of these criterion (pp. 171 - 172), and suggest that Cooper and Holzman apply the stong form of the epiphenomenality "test" detailed by Simon and Ericsson to our (or, indeed any) protocols (p. 172).

Moreover, with particular reference to the studies and model of revision presented here, we must also reject both of the criticisms above. We do not take our data here as "irrefutable proof" of the accuracy or completeness of our revision model. But we do think the model, together with the data, set forth interesting, useful, and testable hypotheses about the kinds of skills experts and novices are likely to differentially exhibit, particularly on controlled revision tasks of the type we have used. We would again like to stress that we are more interested in building a theory of revision processes and showing its plausibility than with testing those predictions which might be implied by this theory. We nevertheless invite research that tests our model against new data obtained in a similar context. And, with Simon and Ericsson (1984), we believe that protocol methodology, if carefully used, would be one valid means of testing some of the hypotheses connected with our model.

(3) A third, final criticism concerns the educational value of expert-novice comparisons (such as the one we are about to report), especially in relation to teaching. In short, why make models of expert-novice differences? What good are these models for improving students' work? To answer these questions, we need to place the present model in the context of our earlier work with experts and novices.

Our previous studies have strongly suggested that writing involves a number of subprocesses whose presence, organization and degree of automaticity (i.e., accessibility to focal attention) will vary with expertise, as well as other factors, such as type of writing task, experimental instruction, and so on. We assume that expertise affects not only what information a subject may heed in short term memory, but also to some extent his ability to report what he heeds. Thus, it

may be that experts, because they have "overlearned" or "automated" a certain subprocess, will not be able to verbalize this process, whereas novices can. For example, an expert might use parallelism in sentences unconsciously and therefore not report information indicating the deliberate choice to use this technique. A novice, however, having just learned what parallelism is, may vocalize information indicating a deliberate, conscious effort. Conversely, both experts and novices might deliberately (consciously) use the same subprocess, but only experts - as a result of better "verbal skills" - might be able to report that they were doing so. Both of these effects of "expertise" are inherent in expert-novice studies of the type we report here, and can probably never be wholly eliminated. In particular, we realize that what may appear to be a difference between experts and novices in underlying writing or revising processes, may sometimes partly be an artifact produced by the interaction between the degree of expertise a subject possesses and the constraint of giving a protocol. But despite these possible confounds in comparing experts to novices, we feel there is nevertheless much benefit in reporting the differences we find, and for two reasons.

First, we assume that some high-level processes (e.g., selection of relevant content while writing or revising) can never be "automated," no matter how "expert" the "expert," and that traces of the information used or produced in the process will be available for reporting in short-term memory. How these higher-level processes are handled seems likely to reflect genuine expert-novice differences, and not merely developmental effects of learning (automaticity) or ancillary effects of verbal "reporting-skill." Second, we not only examine verbal reports but also textual data (in the revisions that subjects actually produce) to support the inferences we are making about likely expert-novice "process" differences. This textual data allows us to check our assumptions about what processes experts and novices appear to employ. In nearly all cases, the process data is paralleled by differences in written output.

As for the "usefulness" of the model itself, we believe it helps us to see more precisely than in any other model produced to date where in the complex process of revision students may experience problems. As Bereiter and Scardamalia (1980) put it, "In order to make advances in teaching a subject like composition, it isn't enough to know what needs to be learned. One has to get to the heart of the difficulties people have in learning it" (p. 3). In particular, the model divides the revision process into several major subprocesses, each of which we describe in detail. We wish to show that any number of different subprocesses can be the source of trouble for novices, and that each of these subprocesses may constitute a skill worthy of separate instructional effort. In other words, to improve students' performance in revision, you may have to improve their ability to perform on a number of discrete subtasks - in addition to increasing what they know about texts, writing and revision.

In summary, we continue to be impressed with the value of expert-novice protocol studies

for theory-building purposes. If better, more "well-defined" theories were available, more theory-testing, using classical psychological performance measures, might be possible. But it is precisely the lack of theory about cognitive processes in writing that motivates our research. With Patrick Suppes (1974) we, too, "want to see a new generation of trained theorists and an equally competent band of experimentalists to surround them, and (we) look for the day when they will show that the theories (we) now cherish were merely humble way stations on the road to the theretical palaces they have constructed."

## Parts Of The Model

We present the model first in our report largely as a means of organizing our observations. Each subprocess is treated in a separate section in our report. The revision model (see Figure 2) is an elaboration of the Review process described in our previous research on writing (see Figure 1). The revison model is divided into two major sections. Processes in which the revisor engages are shown on the left and categories of knowledge which influence these processes or result from their action are shown on the right. Before describing the results of our various protocol analyses, we would first like to present an overview of the different subprocesses represented in the model itself.

Insert figures 1 and 2 here

To perform a task, a person must have a definition of the task to be performed. The **Task Definition** for revision specifies, among other things,

- the goals of the revisor, e.g., revision for clarity or elegance
- the features of the text that should be examined, e.g. global features, local features, or both
- and how the revision process should be carried out, e.g. Should the text be examined in one pass or in several? Should the "gist" of the text be extracted before revision begins?.

We will make two important points about task definition. First, revisors may modify their task definitions during the course of revision. Second, the definition of revision varies from person to person. In particular, experts appear to differ systematically from novices in their goals for revision and in the features of the text they believe should be examined.

The **Goals, Criteria, and Constraints** which define acceptable texts and plans may be ones the writer brings to the revising task or they may be ones that are suggested by needs revealed during the revision process. For example, if revisors notice several wordiness problems in a text, they may set the goal to notice such problems. Goals, criteria and constraints then, may be dynamically modified during revision.

The **Evaluation** process applies the goals and criteria relevant to the revisor's task definition to texts and plans. We believe that evaluation is best viewed as an extension of the familiar process of reading for comprehension. The primary output of evaluation is an initial problem definition, and sometimes important discoveries which influence the revision process. The evaluation process, operating under the goal to modify the text, may thus generate useful alternatives to text features and text goals.

**Problem Representations** vary in specificity or well-definedness. At the ill-defined end of the spectrum are simple detections--the recognition that some sort of problem exists but without a clear specification of the nature of the problem, e.g., "It just doesn't sound quite right here." At the well-defined end of the scale are highly specific problem diagnoses, e.g., "That word is misspelled." Between these two extremes are representations which contain some information about the nature of the problem but still require considerable problem solving activity before the representation is specific enough to allow the revisor to take action on it. Problem representation, then, varies on a continuum for simple detection, through relatively vague diagnoses to very specific diagnoses.

The revisor's **Strategy Selection** depends on the initial problem representation. There are two sorts of strategies available:

1. Those which modify or control the revision process itself: ignoring the problem, delaying action until later, searching for more information to clarify the problem representation.

2. Those which modify the text: revising the text, or rewriting the text.

**Ignoring** the problem occurs when, for instance, a revisor determines that a particular "problem," howvere it is "defined," is not worth fussing over, or, on other occasions, when the definition or nature of the problem itself is not clear to the revisor. **Delaying Action** often occurs when reviewers decide to deal with the text in two passes - one pass for high level problems and one pass for surface problems. On the first pass, the reviewer may notice that the text has spelling problems but decide to wait until the second pass to fix them. **Searching** for more information may occur when the reviewer's diagnosis is not specific enough to suggest a clear choice of action. The search for more specific information has the purpose of moving the representation from the ill-defined toward the well-defined end of the continuum.

When revisors decide to modify the text, they have two major options available to them: **Rewriting and Revising.** By rewriting, we refer to the strategy by which the writer decides to abandon the surface structure of the text, attempts to extract the gist, and rewrites that gist in his or her own words. Rewriting may be done at a relatively local level when the revisor **Paraphrases** individual sentences or at a global level when the revisor **Redrafts** a large section

of the text. Rewriting may be chosen either when the reviser does not have an adequate strategy for fixing the text problem (as is frequently the case for novices) or when the reviser judges that the text has too many problems to make revision worth while (as is often the case with experts).

Because of common usage, we have used "revising" as the name for the whole process by which revisors attempt to improve texts. At this point, however, we want to use the term "revising" in a more restricted sense. By revising in the restricted sense, we understand the strategy in which the writer attempts to fix the text problem while preserving as much of the original text as possible. We assume that successful "revising" requires that the writer can diagnose the text's problems and can apply strategies for fixing them without completely rewriting the text.

Indeed, while this distinction between "revising" and "rewriting" is somewhat difficult to establish, we will present empirical evidence indicating its foundations in the revisors' processes. Specifically, rewriting and revision may be viewed as points on a continuum whose position depends on the extent to which the writer attempts to save the original surface of the text. We assume that writers devotion to preserving the original text can vary continuously over a range of values from high, when one is editing one's boss' memo, to low, when one is repairing the work of a less than competent assistant. In comparison to rewriting, we will attempt to show that revision operates as a focused "repair" procedure based on rich diagnostic information about the problem.

In carrying out the revision process, writers may vary in the goals and criteria they bring to the task, in the kinds of problems they are able to identify in text and in the sophistication of the methods they can bring to bear on their solution. Each revisor may be thought of as having a **Means-Ends Table** in which the problem to be fixed (the ends) are associated with appropriate rules, maxims, and problem solving procedures (the means). As we will see, expert revisors have more detailed and complex means-ends tables than do novice revisors.

To summarize, our purpose in presenting this "preview" of the parts of the model has been to show that revision is indeed a complex process, containing a number of critical sub-processes for investigation. However, before we proceed to a discussion of our findings concerning these subprocesses, we would like to describe the basic components of our experimental task and method. The remainder of the report is structured as follows:

III. Method

IV. Task Definition

V. Evaluation

VI. Problem Representation: Detecting text problems

VII. Problem Representation: Diagnosing text problems

VIII. Strategy Selection.

# III. METHOD

## The Task

Our goal was to create a realistic task that called for a full range of rhetorical and stylistic revisions. Subjects were given a letter written to the C-MU athletic coach by another coach, discussing why women are unnecessarily reluctant to participate in college sports. The subjects were asked to revise this letter into a handout for freshmen women (see appendix). We designed this to be a naturalistic task demanding both high- and low-level revisions. That is, although the information in the letter was appropriate, the task implicitly called for changes in voice, genre, rhetorical stance, perceived audience, and style. In addition, we planted a set of stylistic and rule-governed errors in the text, including problems with spelling, punctuation, grammar, wordiness, and diction (see appendix). The subjects were paid and told that the Writing Center would select the best revision.

## Subjects

The fourteen subjects in the study included 7 students (two of whom were identified by grades and teachers as very capable writers; the others as inexperienced); 3 professional writer/editors who frequently did revision, but typically did not have to diagnose or explain problems in texts; and finally, 4 experienced writing instructors who had also had experience as editors.

## Method and Analysis

The study involved three stages during which the subjects were asked to think aloud as they worked. In Stage 1 they revised the faulty text. Stage 2, immediately afterward, was a Cued Recall session in which subjects were asked to explain the changes they had made. In Stage 3, one week or more later, subjects were given sentences containing the 26 "planted problems" in random order and asked to describe any problems they saw. These three measures allowed us to get converging evidence on what problems people actually saw and to separate what people might "know" on the less demanding Single Sentence task from what they actually used during revision.

The protocols from the three sessions was first parsed to identify instances of reading, proposed changes, and comments. The comments were further parsed into numbered clauses as the basic unit of analysis. The texts were analyzed for revisions. All changes made on the text were also noted on the protocol which allowed us to compare comments in the protocol with actions taken. All comments were then typed on individual 3x5 cards which allowed a variety of blind sortings that are described in the studies which follow.

# IV. DEFINING THE TASK OF REVISION

This first study attempts to define the process of "task definition" as it occurs in our model of revision. The study presents data suggesting important differences in task definition processes between experts and novices, and discusses the implications of these differences for our overall understanding of revision skill.

## Task Definition Vs. Problem Representation

The first process illustrated by our revision model is task definition. In a sense, this process is the foundation for all other revision skills, for it reflects writers' underlying conceptions of what it means to revise or improve a text. That is, the task definition process reflects writers' conceptions of what particular activities are involved in revision and also their conception of how these processes are to be managed. Experimental research on problem-solving tells us that the ability to report or describe our task definitions while we are at work on a problem is limited (Simon & Ericsson, 1984). Unless specifically directed to talk about their concept of a particular type of problem-solving, experimental subjects are not likely to do so to any great extent. Nonetheless, the pervasive effect of apparently "tacit" task definitions upon task performance is noted again and again in research. People appear to create and draw upon knowledge when they anticipate complex tasks.

For this reason, we place the task definition process at the top of our model-not because it is necessarily fixed or unchanging during revision, but rather because of its underlying importance and pervasive influence. Its products are the **goals, constraints** and **criteria** that will guide the overall revision process. Moreover, we distinguish the **task definition** process from the **problem representation** process, also shown in our model. We can characterize the difference as follows: Whereas task definition processes reflect writers' conception of relevant revision criteria, revision activities and how to manage these activities over time (such as when to read "for comprehension," and when to evaluate, problem representation processes reflect writers' understanding of what is wrong with the text or their understanding of what the text needs and how to respond when errors or needs are encountered. We would like to further clarify the differences between these two major processes here, because we feel each contributes independently to revisors' difficulties and successes, and therefore each may be the locus of important (and perhaps teachable) skills.

1) The kind of knowledge that makes up **task definitions** is generally regarded as "metacognitive" knowledge, since a "task definition" implies control over the use of other kinds of knowledge (Bracewell, 1983; Flavell, 1977). In other words, metacognitive knowledge consists

of the things you have to know in order to make use of other kinds of knowledge you have. Like Bracewell (1983), we assume that metacognitive processes are conscious processes (i.e., ones that are deliberately chosen) because knowledge of them can be verbalized. A task definition for revision contains a person's knowledge of what general processes might be used when revising a text, and in what possible orders these processes may be invoked. Let's look at a concrete example.

If I am given a text to revise by an experimenter, I may know that I can either read the whole text through first before making changes, or read one small section at a time and make my changes as I proceed without reading through. Hence, I am making a decision about what general processes to use and in what order to apply them. Or, as another example, I may read completely through the text, then decide to evaluate the text (on a second "pass) for only certain within-sentence problems, such as faulty pronouns. A revisor's task definition controls such decisions about overall goals for the text, the constraints that may limit achieving these goals, and the evaluative criteria to be applied. At any given point I may be more or less conscious of the fact that I am making such choices, but certainly the fact that I may switch from one goal, process or criteria to another indicates that to some extent I can manipulate my definition of the task.

In our description of revision, then, the task definition process creates and stores "metacognitive," or regulatory, knowledge. Moreover, we do not wish to give the impression that task definitions are static, or "fixed" once by the revisor at the beginning of a task and never altered later. Rather, we hypothesize that task definitions are periodically "updated" during the course of revision. As dynamic constructs, they depend partly upon stored conceptions (in long term memory) of what it means to revise, and partly upon what information is provided (continually) by the task environment. By "task environment" we mean to include all those sources of information outside a revisor, including the writer's text with the changes made in it so far (Flower & Hayes, 1981). Whenever a revisor accesses new information from the task environment, the revisor's priorities for executing various sub-processes associated with the task may be altered; thus, the revisor's (current) task definition may change.

2) **Problem representation** differs from task definition in the following way: Though problem representations, like task definitions, involve regulatory knowledge, problem representation is a process that focuses upon some particular difficulty or need occurring in the text. As shown in the model, problem representation refers to the information-processing prompted by some specific unit of text, whether this unit is small or large.

We can further clarify this difference between task definition and problem representation with an extended illustration, as follows.

First, here is what we mean by the "regulatory control" effected by the task definition process. Suppose I am reading a letter my boss has asked me to "improve." When I read the letter for the first time, I am likely to be reading primarily to comprehend it, that is, to build a representation of its meaning-the more so if I have no prior idea what it is about. However, even such a naive reading is still directed by the higher-level goal of "improving" the text-so I am not reading purely to comprehend. Such a shift in goals, between reading purely to comprehend (i.e., for some purpose other than "revision") and reading to revise (or "improve") the text, is controlled by a person's task definition. And task definition controls decisions about other kinds of processes as well.

For instance, in the course of reading with the goal to "improve" the text, my effort to comprehend per se may become "sidetracked." That is, I may encounter many difficulties or needs in the letter, and begin to try to **evaluate** them. In attempting an evaluation (e.g., "Should the text be this way, or some other way?"), in effect I cease to read for "comprehension." I shift processes further, and this shift is controlled by my task definition. Of course, task definitions may also limit in an "a priori" way some of the evaluative procedures I will use. I may, for example, choose to read the first and last line of each paragraph, or skim the text rapidly several times in succession. I may decide to read the text in several discrete passes, using each pass to search for only a few kinds of related problems; or, I may decide to read the entire text in a single pass, evaluating each sentence after I read it for a large range of criteria. A person's task definition will delimit, and monitor, these options.

Finally, task definitions set boundaries to the evaluative criteria which my be applied. For instance, for a certain type of writing, the revisor may opt to ignore spelling errors or, conversely, to search chiefly for such errors. Or, the revisor may look chiefly to see if certain questions are answered (e.g., the journalist's "who, what, where, when and why"); if repetition is justifiably used; if sentences exceed a certain length; or if sentences are written in a certain pattern (e.g., active vs. passive voice). Thus, the task definition process sets up certain criteria, in addition to underlying goals and constraints.

The problem-representation process is quite different, as we can see by returning to the "letter" scenario just described. For example, suppose as I read my boss' letter I come upon a sentence that completely halts my reading-a sentence which I have difficulty understanding and which seems unacceptable. I may represent the sentence's "problem" to myself as one caused by its three embedded dependent clauses, or I may simply say to myself, "It has too many words." One of these two "representations" may be more specific than the other. We refer to the processing called for in making such diagnostic choices the **problem-representation** process. The important point we will make about this process is that the problem representation I have determines the value, or power, of the procedures I will use to correct the problem.

An analogy with manufacturing may provide another way of understanding the difference in our model between **task definition, problem-representation,** and **strategy selection.** In revision, task definition acts much like a quality control *manager* in a complex assembly process. The quality control manager sets the goals, criteria and constraints by which assembly and evaluation of the product proceeds. In particular, the quality control manager "schedules" different inspection processes, determining what inspections are absolutely required for product quality. He constructs a list indicating what general evaluative processes are essential and feasible. Based on this list he may also decide which evaluation routines must come before others, which routines may (but do not necessarily have to) come before others, and which routines may interrupt others. In a similar way, task definition in revision determines (1) the revisor's choice of major revision goals, (2) the availability of evaluative sub-processes needed to meet those goals, and (3) the general "schedule" by which these sub-processes might be carried out.

On the other hand, problem representation is analogous to a quality control *inspector* who actually makes the particular performance evaluations, and assesses (or defines) specific problems with the product. Moreover, just as important interaction occurs between the quality control manager and the inspectors in a manufacturing environment, so we hypothesize that considerable interaction occurs between revisors' task-definition process and their problem-representation process. For example, depending upon the results of diagnostic processes brought to him by the inspector, the quality control manager may change his goals and priorities, and thus call a halt-or temporarily interrupt-other evaluative processes just getting underway. Similarly, task definition during revision may set goals which are periodically modified by information generated later during the problem representation process. We will describe in greater detail how we think this interaction occurs (i.e., in the Discussion section of this first study, and again in the "Strategy Selection" section, VIII.)

Finally, in our model, the strategy selection process might be usefully compared to the *line manager's* function, who works just below the level of the quality control inspector. The line manager makes choices about how to tackle the problems assessed by the inspector, and actually carries out the adjustments indicated. In strategy selection, as we will later discuss, the revisor determines what to do in response to particular text problems. And just as we hypothesize that the problem representation process may change the revisor's task definition, so we also hypothesize that strategy selection may change the goals, criteria and constraints that make up this definition.

### What Are Differences Between Experts' And Novices' Task Definitions?

Having defined the basic differences between task definition and the other major processes in our model, we would now like to describe the specific objectives and method of analysis used in the present study. Our basic question was this: what overall conceptions of revision-or what kinds of "task definitions"-would the early portions of the protocols reveal? This question actually has several parts, as follows.

First, we wanted to see whether experts and novices give themselves the same, or different tasks to perform-that is, do they define the task differently, and if so, how?

Second, we wanted to find out if experts pay more deliberate attention to the construction of their task definitions than do novices. For instance, do experts create revision plans and agendas, or other kinds of process or problem "inventories" that shape their subsequent efforts to improve the text?

Third, we wanted to find out if experts "tailor" their task definition to the occasion more so than novices. For instance, other revision researchers have suggested that most revisors are likely to merely invoke stored, relatively fixed criteria and procedures for guiding the process, rather than adapting their task definition to fit the underlying purpose of the text and the difficulties inherent in that purpose. The implication of text-based studies of revision, in particular, seems to be that few revisors-regardless of whether they are experts or novices-deliberately or consciously attend to the task definition process. For example, Nold (1982) concludes that revisors will not set goals and make "plans" for guiding the revision process unless specifically prompted to do so by teachers or experimenters; instead, revisors will simply invoke static, well-learned sub-routines, much like the college student who, instead of performing the task of comparison and analysis set by essay directions to "Analyze the function of the black veil in Hawthorne's 'The Minister's Black Veil,'" redefines his task as a narrative one, in which he merely recounts the events of the story (Nold, p. 11). Similarly, Nold theorizes that most revisors operate largely out of tacit, stored conceptions of revision; in her view, they do not consciously build a representation of the task's demands, nor do they forecast the goals, criteria and constraints they might use before setting to work.

However, this theory is based upon scant direct evidence. Few (if any) revision studies have attempted to produce evidence of revisor's deliberate attention to the task definition process, and to the "tailoring" of revision goals that may be part of this process. In other words, no revision studies have tried to capture and describe how revisors produce task definitions under experimental conditions. The study reported here is the first to take a close look at how expert and novice revisors initially construct task definitions, and to observe some of the variables apparently involved in this process.

## Method

As noted above, studies of a wide variety of problem-solving tasks indicate that task definitions usually are not verbalized by subjects in detail (Simon & Ericsson, 1984). Indeed, task definition information appears difficult for many people to verbalize concurrently with problem-solving, unless such information can be heeded in short term memory. To deal with this problem, we designed our directions so that revisors could more directly heed the information associated

with task definition. Specifically, while our directions imply a distinction exists between "revising" and "rewriting," the directions themselves do not offer an explicit explanation of this distinction. We deliberately created a sense of dissonance. We wanted subjects to work with the existing text to whatever degree they felt appropriate, so we left it up to them to make their own interpretations of the "revise/rewrite" distinction, with only one constraint imposed: we instructed them not to completely rewrite the text. That is, we did not allow subjects to simply read the target text, set it aside, and then create a new, altogether original text on a separate sheet of paper. We directed subjects to make all their changes on the target text itself, and to use extra paper only when they wished to make additions or do extensive restructuring, etc. Thus, we did not allow them to construct a total, "from scratch" rewrite based upon what they surmised be the purpose of the original. Otherwise, they were free to do as they wished, adding whole new sections if they desired.

Think-aloud protocols of seven experts and seven novices were examined. To determine what differences there might be between the task definitions of experts and novices, we examined all transcripted comments made by both subject groups, beginning with the moment they began reading the task directions and ending when the subjects made their first text change. In particular, we felt that comments in the early stages of the revision task might provide useful evidence of the different features of experts' and novices' task definition process.

To clarify our method, we need to briefly explain exactly how the initial phase of the protocols was segmented.

1. Direction-reading phase. We transcribed subjects' comments during the direction reading phase of the task because we hoped to observe the possible effects these directions had upon the reviewers' initial attempt to build task definitions. The direction-reading phase includes all comments made during the time subjects read the task directions, up to the point where they began reading the target text (i.e., the letter).

2. Initial text-reading phase. This phase is also important for understanding task definition processes. It includes all comments the subjects made between the time they first began reading the target text and the time they made their first text change (i.e., actually added or deleted a word, phrase, and so on). Our major assumption in examining subjects' comments during this phase is that the amount and kind of information-processing revisors do here may have powerful implications for the task definitions they subsequently use to guide task performance.

The comments made by revisors during both direction-reading and text-reading stages were carefully analyzed. Such comments appeared to fall into one of six "activity" categories (Figure 3). These categories were established by an inductive process: that is, all comments made by subjects within the segments described above were sorted into piles, according to the kind of activity the comments seemed to imply. The categories were not generated prior to inspecting the

data. Though the first category, "Reading Completely Through," is probably self-explanatory, the others may not be, and will be briefly explicated here.

(Insert Figure 3 about here)

Category #2 refers to comments in which the subject introduces either a "plan" or "goal" for the task that was not explicitly provided by the task directions, and that was not a mere paraphrase of these directions. For example, a subject might have said, "So my objective here is to really pump up the enthusiasm of freshman women for sports." Category #3 refers to comments in which the subject summarizes the contents of the letter or its purpose, as in, "So basically this handout just compares advantages of intramurals with something else you might do in your leisure time at school." Category #4 refers to comments in which the subject focuses upon the audience's possible needs-as in, "I think this kind of document ought to be simple, because Freshman students have very short attention spans!" Category #5 refers to comments in which subjects build a list or schedule of things they think will need to be done. For example, "The first thing I'll do is remove all the spelling errors, then the vague pronouns. Then I'll try to get the larger organizational matters taken care of." Category #6 refers to comments in which the subject mentions some negative impression created by the text as a whole, or some negative trait that can't be localized. The subject might comment, for instance, "Somehow the voice in this thing is wrong" or "It's really boring." Finally, Category #7 refers to comments in which the subject identifies some particular fault that must be corrected, though the subject does not immediately proceed with correction, as in: "Those exclamation points have to go," or "That's a vague pronoun if ever I saw one."

The reliability of these categories was checked by four independent coders, each of whom coded 26 comments in an actual protocol. Average inter-coder reliability was 81%.

## Results

Interestingly, two of the novices and one of the experts (a professional writer) did not read the target text through completely before beginning to make changes. Instead, they read the first few sentences of this text, and then immediately began adding and/or deleting portions, without reading ahead to the end. Obviously, revisors who do not read completely through the text at least once do not draw upon the same amount of information in constructing their task definitions as those who do. Accordingly, one might assume their task definitions will be less appropriate than those of subjects who do read completely through. Not reading the text through might be taken as an indication of poor revision skill, i.e., a failure to attend to all the necessary information. A possible explanation as to why some subjects did not read completely through in advance is simply that they might find the task so familiar or routine that reading all the way through did not seem important to them. Thus, on the face of it, such behavior should not necessarily be construed as a sign of poor revision skill, or even as a handicap on performance, without taking

other data into consideration. This behavior does indicate, however, that incomplete-readers are likely to form a rather different task definition than are complete-readers.

In Figure 3, an "X" indicates that a subject made at least one comment in a given category; a "O" indicates that a subject made no comments in a given category. The first two columns in Figure 3 show the number of subjects from each group making one or more comments in a specific category. As the Figure indicates, experts scored significantly more than the novices when the six comment categories are aggregated. Experts scored in 20 of the 42 cells, compared to the novices' score of 6 out of 42 cells, or 47.6% (experts) vs. 14.4% (novices). These results were significant using a Mann-Whitney U-Test ($p < .03$).

**Discussion**

We do not assert that comments in these categories constitute unequivocal evidence of specific processes occurring during task definition. However, the categories do point to metacognitive skills that we feel are worthy of further investigation. For by the time expert revisors make their first text change, our data suggest that they have created richer, more complex task definitions than novices. Experts are much more attentive to the rhetorical context of the document, and to the "global" problems associated with this context. Indeed, experts not only have processed different information about the task, establishing markedly different revision goals and plans, but also they apparently have more regulatory processes available for storing and retrieving information during subsequent efforts to improve the text. In particular, the evidence supports the following observations:

1. *Experts have more knowledge about how to make process plans to guide task performance (Category 2).* For example, one expert (HD) during his initial reading of the target text commented that "one of the things that I do normally day to day is, when I first look at a piece I very quickly see the spelling errors, and I tend to make the changes in that *in my first run through* because sometimes as you read through it later you may miss them . . ." Another expert's comments (JP), upon reading the text through once, indicate his awareness that revision is a complex task whose subtasks need to be regulated; he asks rhetorically, "Do I describe *each revising pass* as I do it?" Only one of the novices made any comments suggesting knowledge of process plans.

2. *Experts deliberately create an "inventory" of problems in anticipation of subsequent work (Category 5).* For example, an expert (JP) revealed that he tries initially to organize the "problems" he will focus upon in a kind of list-a list that may also serve as a means of managing the flow of his attention: "the first thing I will do at least as I read the sentences [again], I mean the thing . . . it seems to me the person does have a sense of organization, does write what I call topic sentences, and sort of . . . there's just plain errors, and there's wordiness in the prose . . . well, I have two basic biases in revising, or three. One of course is mechanics . . . one is cutting out as many adjectives as don't seem necessary, and the [other] is saying things like 'assume' instead of 'possess the assumption that' . . . I know I'm going to come back-want to come back and do something with what seems to be a bad voice." Another expert (HD), also constructs an initial inventory

of problems as a way of managing his attention. He remarks, "The thing that jumps out at me here, you know this is a, one of those kinds of letters, the exclamation points and that kind of 'up' thing. Maybe it's my personality, but I'm not into that kind of thing. I think where I find exclamation points I'll quickly dump them... I don't like overstatement, and there's some of that in here . . ." None of the novices made comments suggesting that they built "inventories" or schedules before actually making text changes.

3. *Experts' task definitions include "global" goals for revision that take into account the purpose of the text in the communication situation (Category 6).* One expert subject (BA), for example, remarked in the course of her first reading that "It seems to me that if the goal of this article is to get women involved, it ought to be clear whether they want them to be involved in intramural or varsity sports, or both . . . right at the beginning." By contrast, only two of the novices made any comments that suggested awareness of "global" problems with the text before making text changes.

While data for each category of results (1 through 6), taken separately, is insufficient for determining statistical significance, the overall pattern of results between the two groups suggests that experts build strikingly different task definitions than novices and also bring greater "managerial" skill to performance. (This "managerial skill" is further evidenced by choices in revision "strategy selection," a process we discuss in detail later).

First of all, when the data for categories 2, 5 and 6 are taken together, it becomes strikingly clear that from the outset experts are building completely different task definitions than novices. Experts make an effort to "tailor" their conception of revision to the task at hand. Ultimately, experts are doing a different task, focusing much more globally upon the purpose of the text in the particular communication situation. By contrast, novices appear to conceive of revision in a radically different way, by searching for problems in the text without much attention to the text's purpose or to global aspects of the communication situation. The evidence further suggests that they are not attending to their task definitions and therefore not actively constructing or "tailoring" them before they revise. Indeed, the evidence from the initial segments of the protocols leads us to infer that they do not set goals or make plans for approaching the task in a comprehensive way. Rather they seem much more directly dependent upon the order in which problems appear in the target text. They do not revise on the basis of problem "inventories" or similar ad hoc "schedules" for regulating subprocesses. Thus, whereas many experts in the initial stages of their protocols did appear to "tailor" their task definitions as they read, most novices seemed to apply a fixed, stored routine.

Second, the results suggest that experts may see revision as a collection of more or less distinguishable processes which must be approached iteratively, in successive "passes." Indeed, protocol comments suggest that experts are more aware than novices of the complex demands which revision places on attentional capacity, and as a result consciously try to break down

revision processes into more discrete subprocesses. This inference, at least, is supported by the results for Categories 2 and 5, where experts "forecast" the kind of criteria they may employ over the course of discrete subsequent "passes." Some experts explicitly set about evaluating text for certain kinds of problems before others, thus recognizing a principle of economy of effort, or what we call "precedence guidelines." The experts appear aware that revising involves applying many different kinds of procedures over time, and that determining in advance a rough order for these procedures is likely to make them more efficient.

To summarize, key differences in revision behavior appear in the way experts represent their goals for the task from the outset, in the way they represent the structure of the process to themselves (i.e., single pass with no priorities vs. many passes with some priorities set) and also in the way experts manage the processes associated with the task. Given these findings, we think that there are several skills associated with task definition that might be addressed independently: (1) novices not only need to learn what information to attend to in the text they revise, but perhaps more importantly (2) how to set goals and make plans for revising, and (3) how to support (or "monitor") their attentional capacities while work is under way. Such "expert" control methods as informal inventories, or plans that assign different subtasks to different "passes," might reduce information overload, and thus make revision more efficient. In particular, to build these "inventories" and set priorities (or hierarchies) within them means that the novice revisor also needs to learn (4) to attend to the underlying purpose of the communication and the problems inherent in this purpose. Without such attention, it would seem difficult (if not impossible) for a revisor to construct meaningful revision plans, or to monitor progress in carrying them out. Where such attention to purpose is lacking, the revisor is likely to pick out problems more haphazardly.

On the basis of this first study, then, we believe several issues deserve more investigation. First, we need to investigate the reasonable assumption that revisors' task definitions, and in particular the revisor's goals for revision, vary in response to a host of factors. For while the experts as a group clearly created different initial task definitions and goals than the novices, there was nevertheless a wide range of initial definitions and goals among the experts themselves. These differences are indicated by the fact that individual experts did not have the same pattern of comments in the categories (Figure 3). The factors affecting initial task definition and goals include (at least) the following:

1. attention to the purpose, length and genre of text to be revised
2. attention to the explicit verbal demands or directions presented to the revisor by an instructor, experimenter or other instigator
3. attention to the implicit demands represented by the intended audience and by the context this audience shares with the revisor

4. attention to stored conceptions of the revision-process and to textual knowledge in the revisor's long-term memory

5. attention to changes introduced to the text as the revisor works at the task

6. attention to the number , density and complexity of problems and errors in the target text

The two initial segments of the protocols we examined in this study showed that attention to all of these factors influenced task definitions. And because this list of factors is probably incomplete, we feel safe in concluding that there is no such thing as an "optimal" task definition for any particular revising task. Depending upon how many and which of these factors are attended to, a number of different, but equally appropriate task definitions might result. However, by experimentally manipulating the amount of attention the revisor may place upon one or more of these factors, we might be able to discover more about how each affects the task definition process and the goals that revisors initially establish. For instance, here is how we might examine the effect of (1) error number, (2) error complexity and (3) error "density" in texts (item #6 in the list above) upon task definition: texts varying along these three dimensions could be submitted to revisors, to see if there is a "saturation point" at which both novices and experts alike will opt to completely "rewrite" a text instead of "revising" it piecemeal. Indeed, some of our experts commented that they wanted to simply "rewrite" the text because of the large number and complexity of problems they found. One subject (DS) commented, "I wouldn't even write it this way. I'm just so. . . This thing is so ridiculous. . . I want to rewrite it. . ."

Another study might focus upon the effects of "problem inventories" upon attentional capacity and "managerial" control of the revision process. For instance, novices might be trained to read completely through stimulus texts and to make prioritized inventories of the problems they find as they proceed. Their control over the revision process, along with the effectiveness of their final revisions, could then be compared with that of a group of (matched) novices who attempt the task without such training in "inventory making." To establish the precise contribution of the inventory method, both groups could be carefully instructed in methods for solving the types of problems presented (e.g., faulty pronoun reference, faulty parallelism, etc.)

As noted earlier, we hypothesize that task definitions are not static but rather dynamic entities. Task definitions appear to change as work on the revision task progresses, and we belive our results here justify study of how and why they change. Specifically, our protocols suggested that experts' task definitions and problem representations will interact more than those of novices over the course of the whole revision process. We can explain what we mean by "interact" by referring again to our model and our earlier hypothetical example. Suppose I have chosen to read my boss' letter and revise it as I proceed. That is, I do not read all the way through the

text before beginning to make changes. Now, as I am reading and checking, suppose that I encounter ("detect") an unacceptable sentence and I pause to represent its problem as being triple dependent clauses. I decide to remove immediately two of these clauses to solve the problem. However, in so doing, I reconsider whether I ought to go ahead and simply read through the rest of the text without stopping, merely flagging other sentences with this or a similar problem. After pondering a moment, I do decide to finish reading the text and to postpone further correction, merely flagging other instances of the same type of problem.

Notice what has happened in this scenario. My "detection" and "problem-representation" have altered my initial task definition and thus changed my process plan for the task. In particular, I have "updated" my initial (or "working") task definition, which directed me simply to revise as I read. At this point my task definition changes because I have now decided to read all the way through, paying special attention to dependent clauses as I go.

We feel this interaction, or "feedback loop," between the problem representation and the task definition, is very significant. For it may be that experts are more able than novices to adjust their task definition in response to specific problems detected in the course of revising. In other words, the experts will continue to adjust their task definition and process plans in response to new information, whereas novices may continue to inappropriately "force" an initial, "low level" task definition to do the job, i.e., not making later adjustments in their revision process plans. Though our evidence in this first study does not specifically address this point, our findings here (and in the other studies in this group) lead us to hypothesize such a difference between expert and novice performance. In other words, we predict the task definitions of experts will be both more "tailored" in the initial phases of the task and more "flexible" over the course of performance.

Our data also indicate that the reverse process is a significant one for distinguishing experts and novices. For instance, the data in this study strongly implies that the information in the "initial" task definition (i.e., "goals and criteria") subsequently influences both the detection and representation of problems in the text. To express this phenomena in a metaphor, the task definition acts like a pair of colored glasses: if I have a task definition that considers revision as primarily the deletion of excess words, then I may not detect other sorts of problems in the text (e.g., lack of specificity in illustrative examples). I simply won't see certain kinds of difficulties. Thus, the task definition may force some problems to go undetected, or, what is equally interesting, it may influence the "strength" or "definedness" of diagnoses that revisors make of specific problems they do detect.

Indeed, later studies in this series provide evidence that inappropriately narrow task definitions may be the culprits that prevent novices from making "detections," independently of a

lack of specific knowledge about the problems themselves. That is, unless a more global task definition is tailored for the occasion by the revisor, even problems for which the novice has the requisite diagnostic knowledge may go unsolved. In the model, we show this influence of task definition upon problem representation with an arrow from the revisor's "goals and criteria" for text to the "evaluate" function (Figure 2).

# V. EVALUATION

In this section, we will present a more detailed model of the Evaluation subprocess and suggest its relation to the process of reading. We believe that evaluation is best viewed as an extension of the familiar process of reading for comprehension. We will first discuss reading for comprehension and then describe how this process is modified to Evaluate and Define problems in text.

## Reading For Comprehension

By reading for comprehension, we understand a process by which the reader attempts to construct a satisfactory internal representation of the meaning of the text. Consistent with Thibadeau, Carpenter and Just's (1982) model of reading, we assume that the process by which a reader constructs an internal representation of the text involves the interaction of subprocesses on many levels. Our proposal for such a multilevel representation process is shown in Figure 4. The processes we have chosen to include in the model differ from those in the Thibadeau, Carpenter and Just model particularly on the more complex levels. Our choice of processes is based on our analysis of thinking aloud protocols of subjects reading and attempting to understand fairly difficult technical prose (See Schriver, forthcoming). The levels of processing include at least the following:

1. Decoding words, that is, identifying individual words and retrieving their meanings from memory.
2. Identifying the grammatical structure of sentences.
3. Applying elementary semantic rules, e.g., interpreting sentences such as "The pen is in the coat." and "The coat is in the pen."
4. Making factual inferences.
5. Applying schemas and world knowledge.
6. Making use of genre conventions, e.g., "I guess he is summing up here."
7. Identifying gist
8. Inferring the text's function or the writer's intentions and point of view.

insert figure 4 here

We are not asserting that these subprocesses act in sequence. Rather, we assume that they may occur in any order and that they cooperate with one another in the effort to produce a satisfactory text representation. For example, finding the grammatical structure of ambiguous sentences such as "They are sinking ships." may depend on identifying the gist of the text.

Further, we assume that all of these subprocesses may make use of all of the information about the text that is available at the time--that is, that readers integrate their interpretations of the new text with the interpretation they have made of the text so far. Thus, decisions about newly read text concerning the meanings of words, grammatical structure, the writer's plans and goals, etc. are all made so as to produce an integrated representation of the text and of the author's intentions.

Reading for comprehension, by itself, can lead to detection of some problems under the following circumstances. When everything goes smoothly, as it usually does in reading well constructed text, the reader's attention is focused primarily on the representation being produced and very little on the comprehension process that is producing it. With poorly constructed texts, however, any of the subprocesses may encounter comprehension difficulties which bring text problems to the reader's attention. Predictably, when people are reading for comprehension, they appear not to be nearly as sensitive to text problems as they are when they are revising. However, reading protocols indicate that even while reading for comprehension, readers do retain sensitivity to some kinds of text problems. For example, if the reader sees "resive" in the text, decoding this string as the word "receive" may require some conscious problem solving. As a byproduct of this effort, the reader will likely be aware that there has been a misspelling. Similarly, the reader may become aware of other text faults such as grammatical problems and errors of fact. For some problems, the editor never sleeps. The right hand side of Figure 4 shows some problem categories to which readers appear to retain sensitivity. The arrows in the diagram are intended to indicate that the detection and diagnosis of different text problems is associated with different levels of the comprehension process. That is, the detection/diagnosis of spelling faults typically is the result of the word decoding process and the detection of incoherence, of the process of identifying gist.

### Reading to Evaluate and/or to Define Problems

People may read a text with a variety of purposes. They may read to comprehend. They may read simply to evaluate, e.g. "Is this essay sufficiently well written to be included in my anthology?" And they may evaluate in order to revise. When writers read a text to revise it, they still read to comprehend it, but they adopt several goals in addition. First, they read with the goal of detecting and characterizing text problems. This goal makes them responsive to a much wider variety of text problems than people reading simply for comprehension (Compare the right hand sides of Figures 4 and 5). Second, they adopt the goal of fixing the text problems that they find. This goal leads them to take a more active attitude toward the text, searching for ways to improve it at all levels from wording to overall text structure. Searching for possible improvements may lead to discoveries associated with various levels of the comprehension process as shown in the left hand column of Figure 5. For example, searching for an alternate word in order to avoid repetition may trigger an association to another part of the text and reveal a

surprising relation within the text. Or searching for an example to illustrate a principle may lead to the discovery of a way to make contact with the special interests of the audience. Third, if the revisors are skilled, they adopt the goal of shaping the text to the needs of the intended audience. This goal has a profound effect on performance at all levels. For example, it leads skilled revisors to consider not only whether the words of the text are intelligible to the themselves but also whether they are intelligible and acceptable to the intended audience.

Insert figure 5 here

These additional goals make the behavior of the revisor quite different from that of the person reading for comprehension. Revisors are more concerned with detecting and diagnosing text problems, are more actively searching for alternative text features and are therefore more likely to make useful discoveries about the text than are those reading for comprehension.

The three activities shown as part of the evaluation process in Figure 2--Reading to Comprehend, Reading to Evaluate, and Defining the Problem--can be distinguished by a number of measures: 1) They produce different results (ranging from automatic detections, to evaluative judgements, to elaborated problem diagnoses). 2) They often take different amounts of time (e.g. defining the problem often requires considerable problem solving effort beyond that entailed in reading to evaluate or simply to understand). 3) Finally, they are associated with expert/novice differences. Here again, defining the problem is normally a demanding process that novices rarely display.

We see the sequence in the evaluation process in Figure 2 as reflecting an escalation of goals and constraints from reading for comprehension as the least constrained to reading to define problems as the most constrained.

While revision may involve all of the levels we have specified in Figure 5, novices appear to entertain fewer constraints and they typically make less use of the three most complex levels than do experts. Novices are less likely than experts to infer gist, to infer the writer's intentions, and to consider the reader's needs.

## Special Revising Situations

To discuss the revisor's performance in more detail, we must distinguish between three rather different evaluation situations:

1. Revisors evaluating another writer's text.
2. Revisors evaluating their own text.
3. Revisors evaluating plans whether their own or those of another writer.

Revisors evaluating another writer's text.

When revisors are evaluating text written by other writers, they have only indirect access through the text itself to the writer's plans which underlie the text. In many cases, the evidence in the text is clear enough that the revisor can be confident of the writer's original intent. In some cases, though, the evidence in the text is not clear. For example, one of our revisors said of the text, "I hope that's what she meant to say because that's what she did say."

Revisors evaluating their own text.

Revisors who are revising other writers' texts can only reconstruct the writers' intentions from the text itself. Revisors who are revising their own texts have privileged access to the intentions which gave rise to the text. This difference in access to the writers intentions can be both a help and a hindrance. It can be a help because in many cases, only the writers knows if a given word or phrase accurately matches their intention. It can be a hindrance when, as Bartlett (1980) has shown, writers' knowledge of their own intentions may blind them to certain faults in their texts.

Writers revising plans.

According to the model shown in Figure 2, when revisors evaluate a text, they may also evaluate the text plan as a part of the text evaluation activity. We distinguish the evaluation of plans from the evaluation of text at this point for three reasons. First, plans are sometimes evaluated separately from text. A writer may construct a plan for writing a text and then evaluate and revise that plan before any text has actually been written. Second, the process of evaluating a plan is somewhat different from the process of evaluating a text. Evaluating a plan typically does not involve consideration of the surface structure of the text that the plan would produce. Indeed, writers frequently find that plans that appeared quite adequate simply don't work when one attempts to write them down because there is no acceptable surface structure to embody the plan. A common example is attempting to express too many ideas in one sentence. We simply don't recognize the weakness of the plan until we try to execute it. Finally, we want to emphasize the importance of the evaluation of plans in the revision process whether the writer is dealing with a text or not. While, as we will show, experts are superior to novices in detecting and diagnosing surface level faults in a text, the weight of evidence suggests that the real advantage of experts lies in the greater attention they devote to the evaluation and revision of text plans.

## VI. PROBLEM REPRESENTATION: DETECTING PROBLEMS IN TEXT

Anyone who has tried to advise a friend (or oneself) on a poorly written draft is likely to say that reading even a hard text is far easier than trying to diagnose or revise it. Perhaps it is this intuition that has led most revision theorists to treat evaluation as an isolated act of comparison and testing. In contrast, one of the chief suggestions made in Section V was that evaluation is best understood as an extension of the reading process. The sub-processes we identified within Evaluation represent distinct locations on a continuum marked by an increasing number of goals and constraints imposed upon one's representation of the text. As reader/revisors move along this continuum, they metaphorically turn up the power or work load of the process--subjecting their growing representations of the text to more stringent demands and attempting to define its failures more fully.

The output of the Evaluation process--at whatever level it is conducted--is a Problem Representation (see Figure 2). Problem representations also vary along a continuum for which the metric is not power or effort but the amount of information the Evaluation process has produced. In discussing Evaluation, then, we focused on the conduct of a process. Now we turn to its output--the writer's problem representation--as it is reflected in our data from expert and novice revisors.

The process of revision simply can't proceed until the revisor perceives that the text, or a plan for producing a text, has problems. Problem detection, then, is a necessary precondition for revision. In this section, we will describe two differences between college students and experts in detecting text problems: First, college students persistently fail to detect some problems which experts detect easily. Second, experts are much more likely to detect global problems than are college students.

As noted above, our experimental procedures involved the collection of thinking aloud protocols from subjects as they revised a text with 26 planted errors and a number of global problems. The plants included misspellings, faulty parallelism, etc. Most of these plants were problems at the sentence level or below. The global problems included such things as inappropriate focus on the audience and improperly sequenced topics. The protocols allowed us to determine when the revisors were reading each text problem and exactly what their responses to the problem were as they examined the text.

The novice revisors typically read each problem more than once. In some cases, they read particular problems as many as eight times. Their responses included detection of the problem,

failure to detect the problem without comment, and failure to detect the problem accompanied by comments indicating approval of the text containing the problem. The novice revisor, GB, persistently failed to detect text problems in the following sentence while reading it on four separate occasions: "In sports like fencing for a long time many of our varsity team members had no previous experience anyway". After his first reading, GB said "Ok". After the second, he made no direct comment. After the third reading he said, "Ok, that sounds good right there." and after the fourth, "Ok." At no time did he indicate in any way that he found fault with the sentence, and, in fact, on his third reading he appeared to approve of it. Another novice revisor commenting on this sentence said, "Freshmen would like that".

The novices' insensitivity to text problems appears in some cases to be relative rather than absolute. For example, the subject, GB, repeatedly read the sentence which began "Many naive women students possess the assumption that. . ." On the first reading, he deleted "naive" but made no other changes. On the second reading, he said "Ok, that's good. . ." Finally, on the third reading he said, "No, no, no. . .shorten it up." and changed "possess the assumption" to "think." Although some of the difference in sensitivity is relative, some is, in the practical sense, absolute. That is, there are some text errors which the novices never caught despite repeated rereadings of the text. Thus, another novice read this same sentence seven times and never found fault with it.

Figure 6 shows the total number of times novice subjects read text segments containing errors, the percent of these occasions on which they detected the errors (26%), the percent of occasions on which they actively expressed approval of the flawed text (6%), and the percent of occasions on which they read the errors without either comment or action (68%). Adding the last two categories together, we find that novices had a 74% failure rate in detecting the problems they read.

insert Figure 6 here

We can't provide a comparable summary table for the experts because, as will be discussed below, many of the experts were playing a different game than the novices. That is, their strategy of revising did not require them systematically to comb the text in order to detect problems. Rather, some decided that the text had too many errors to make it worthwhile repairing each one individually. They decided, therefore, simply to extract the gist of the text, to rewrite it in their own words, and to ignore the individual errors. As a point of comparison, the expert whose revision strategy most closely resembled that of the novices had a failure rate of 36% in comparison to 74% for the novices.

It is conceivable that, despite the lower efficiency of their detection processes, novices might, through persistence, find as many of the errors as the experts do. However, as Figure 7

shows, they did not. Evidence from the protocols indicated that by the end of the revision sessions, the experts had detected 58% of the planted problems while the students had detected only 36%.

insert figure 7 here

When we examine the percent of problems eliminated in the revisions, we notice an anomaly. The experts eliminated 91% of the errors and the novices eliminated 64%. In each case, these figures are about 30% higher than the percentage of problems detected. How can we account for this anomaly?

There are, it seems to us, two possible reasons why revisors might fix or eliminate apparently undetected problems:

1. Some problem detections are missed in the protocols. Thus, some apparently undetected problems that were fixed were in fact detected by the revisors but not mentioned in the protocols.

2. Problems get eliminated which, in fact, were not detected. This might happen in two ways:

   a. Revisors who adopt a rewriting strategy extract the gist of text without attending to its errors. They then generate new text on the basis of their own production grammars which are relatively error free.

   b. Revisors may accidentally eliminate a problem they have not detected while attempting to fix other problems which they have detected.

Our protocol evidence indicates clearly that the second hypothesis is correct. Deciding whether or not the first hypothesis is also correct will require further experimental work.

Examining the protocols also makes it clear that there was a difference not only in the number of errors the experts detected but also in the kinds of errors they detected. Not only do the experts detect more sentence level problems--the kind of problem to which novices devote almost all of their attention--but more global problems as well--a kind of problem novices rarely detect.

With one exception, the planted problems were strictly sentence level problems. The global problems that the experts detected, e.g. the recognition that a sentence was placed in the wrong paragraph, were, in most cases, additional to those counted in Figure 7. Thus, the data in Figure 7 underestimates the difference in problem detection between expert and novice because it focuses primarily on sentence level problems. The protocols suggest strongly that the experts were more concerned than the students with global problems.

To provide quantitative confirmation for this hypothesis, we selected two expert and two

novice protocols at random and analyzed them in the following way: First, the protocols were divided into problem solving episodes. (Judges can identify such episodes quite reliably.) Typically, episodes correspond to a single problem solving event, e.g., "Ok, what am I going to do about this monster? I'm going to delete it, that's what."

After the protocols were divided into episodes, each episode was evaluated to to determine whether it was concerned with a problem above the sentence level or a problem at or below the sentence level. Figure 8 shows the proportion of episodes in each category for both experts and students. For the novices, only 15% of episodes involved problems above the sentence level in contrast to 42%, or nearly three times as many, for the experts.

insert Figure 8 here

## VII. PROBLEM REPRESENTATION: DIAGNOSING PROBLEMS IN TEXT

We have noted that people detect many problems simply because their own reading process is derailed by the text. But when people are reading to Evaluate, their sensitivity to text problems is heightened and directed in specific ways. We have also noted that Evaluation per se is a constructive process--just as reading builds a representation of a text, Evaluation builds a representation of problems. In this section, we will look more closely at the nature and content of these representations. First, we will describe our model of the continuum along which revisors' problem representations exist; we will also explain what we feel are the conceptual advantages of using this continuum to characterize problem representations. Second, we will explain the significance of a critical segment (or "range") of the continuum - a segmant comprised of what we call "diagnoses." Third, we will discuss a variety of findings in our protocols that helped us to generate our continuum model of problem representations.

### The Continuum of Problem Representations

To begin with, the problem representations people actually create appear to exist along a continuum. Representations range from spare representations which contain little information about the problem to richly elaborated diagnoses which offer both conceptual and procedural information about the problem. Problem representations found at the low end of this continuum may be expressed by such protocol comments as "I don't like this." Or, when asked why a change was made, the writer may simply assert, "I don't know, it just didn't sound right. In Figure 9 we have labeled such representations as "Simple Detects." They contain little information beyond the recognition that a problem exists. Such representations are simple, efficient and often sufficient for the task, but they also indicate that the writer is working with limited information and an unelaborated representation of the problem. He has given himself an ill-defined problem to solve.

The continuum in Figure 9 then is based on the amount and specificity of information contained in a revisor's representation. (Note that this continuum is not a quality scale based on an assumption that some representations are inherently better than others.) At the far right we find those representations which contain a good deal of highly specified information about the problem and, as a consequence, give the writer a well-defined problem with which to work. Such well-defined problems are most frequently associated with violations of a spelling or grammatical rule as in the sentence, "athletics is an important aspect of college." When the revisor represents the problem to himself as "a disagreement between the subject and verb," he

has pinpointed the location of the problem (to the two words "athletics" and "is"), has identified the cause (disagreement in number) and in the process has specified a clear solution procedure (change the verb). The subject whose "simple detect" represents the problem as "this sentence just doesn't sound right," has given given herself a substantially different problem.

(Insert Figure 9 here)

Unfortunately for writers, most of the significant problems in adult texts can not be turned into completely well-defined problems and are located by necessity somewhere short of the far right of this continuum. For example, matters of diction and style (e.g., whether to use a technical term or when to use "I") are not governed by rules, but by judgment, taste, and a sense of appropriateness. More global problems such as voice, persuasiveness, or organization are by nature even less well defined. They are harder to attach to a specific word, their effect is more diffuse, and there are fewer explicit procedures for dealing with them. Revision, unlike simple proofreading, is the process of solving a series of interlinked, but relatively ill-defined problems.

Two variables appear to affect where a given representation fits on this continuum. One is the nature of the problem itself, as discussed above. The other is how fully the writer chooses (or is able) to represent the problem. For example, a perception of incoherence in a paragraph could be represented with a "simple detect" ("wait, this is confusing") or with additional information about location ("these two sentences don't work"), or with a general rhetorical concept such as "flow" or with a more technical concept such as "topic shift" which contains procedures for tracing and testing the topic development and heuristics for repairing shifts.

As our model emphasizes, problem representations range along a continuum from ill- to well-definedness. However, striking substantive differences in some of these representations led us to identify a special class of high-information representations, for which we have reserved the term Diagnosis. A <u>bona fide</u> Diagnosis, unlike other forms of representation, depends on concept recognition. When diagnosing, the revisor recognizes the current problem as belonging to a class or category of text problems the revisor already knows. For instance, people often detect that they feel peaked and have a fair amount of information about their condition, but still do not know what the problem is. That's when a doctor's representation is called for. So by the same analogy, a diagnosis is a problem representation you would be willing to pay for. When a revisor says, "that's terribly wordy" or "this never comes to a point," or "I don't see the logical connection between x and y," she places the problem in a conceptual framework which brings with it information about typical problem features, tests for refining the definition, and likely procedures for fixing the problem.

Diagnosis depends on concepts and categories. But note that it does not depend on names.

A revisor does not need to know a technical term or even a name at all in order to diagnose a current problem as yet another "whatchamacallit." The critical feature of a diagnostic concept is, of course, not its name, but the useful prior knowledge about texts and discourse which it brings with it.

## Conceptual Advantages of A Continuum Model

Defining problem representation as an ill- to well-defined continuum with the special category of Diagnoses at the upper end has certain advantages for building a theory of revision processes. Such a continuum allows us, for instance, to account for some of the limitations of novice performance. Scardamalia and Bereiter (1983) found that 6th and 8th grade students could identify certain problems such as, "people may not believe this" or "I think this could be said more clearly" with some reliability (they agreed with an adult). And yet, when asked to explain or support why they had chosen one of these evaluative statements from a set of cards, they could go no further--they made uninterpretable comments or simply reiterated the evaluative card. Why? One explanation is that these students, with the aid of their evaluative cards, are working rather effectively at the level of mid-range detection. Their problem representations can indeed go beyond the level of "simple detects": they include information about the revisor's reading experience, the general location of a problem, and social knowledge about how readers respond to topics. However, they are not able to diagnose the problem more fully or move up to a categorical representation, much less one that involves specific rhetorical knowledge about discourse patterns and text.

A second conceptual advantage of this continuum is that it helps us to account for the behavior of expert writers who frequently go through a series of problem representations. Although experts may begin like anyone else, with a simple detection, they often exercise the search option (see Figure 2) to loop back into Problem Representation. By reconsidering the problem, they typically move their representation further and further to the right, turning a detection into a more fully defined and more manageable problem.

## The Range of Diagnoses

The previous section has pinpointed the key difference between detecting a problem and diagnosing a problem: in order to diagnose, the writer develops a relatively well-defined representation of a text problem which is based on a categorization of the problem-type. However, diagnoses themselves vary in their level of specificity and in the kinds of problem they define. At one extreme (at the left of our continuum in Figure 9) we find relatively vague and ill-defined diagnoses. One form these diagnoses take is an attempt to simulate a reader's response, as in the following examples.

- I think I'd just read that first sentence and say "why bother"

- No-one would believe that.

In these verbalizations the category of problem is implicit; it may not have a name. However, the revisor has clearly recognized a problem type. Some problems seem to demand these ill-defined diagnoses and cannot be "reduced" to more well-defined ones. Other problems, however, admit very specific, well-defined diagnoses, such as diagnoses of mechanical errors which exhibit an explicit and recognizable category (e.g., ambiguous relative pronoun reference).

In analyzing the diagnostic comments of our expert and novice revisors we were able to distinguish three overall types of diagnosis:

1. The least well-defined diagnoses we classified as "intentional" diagnoses. Here the revisor attempts, with whatever rhetorical knowledge she has at her disposal, to compare her representation of the text's purpose and/or the author's intentions to what is actually written, and then to decide how to change the text (or a text plan) to meet these apparent intentions. An example of such an "intentional" diagnosis might be: "The purpose of this thing is to get women into a large active group to lobby for funding. And the way to do this is to recruit for teams". From this statement, this revisor proceeded to plan possible ways to change the text to meet its apparent intended purpose.

2. The most well-defined diagnoses we classified as "rule-based." Here the revisor has a ready-made problem representation based on her knowledge of violations of standard rules of grammar, punctuation and spelling. She also has a built-in action--if a spelling is wrong, fix it.

3. A number of statements, which fell between these two extremes, were classified as "maxim-based." Here the revisor has some established guidelines, but not any clear-cut rules upon which to base her diagnoses. Such maxims incorporate those found in standard handbooks, such as avoid wordy prose or use parallel structure. They may also be personal to the writer--either general rhetorical guidelines she always employs (such as avoiding negative constructions) or more genre-specific guidelines she has developed, through experience with, for example, writing public-relations handouts.

## Results of Protocol Analysis: Differences Between Experts and Novices

The results in Figure 7 show that experts and novices differed not only in their detection rate but in the number of problems they diagnosed and eliminated. If we consider the 26 "planted" problems (which ranged from spelling, to diction, to sentence style to organization) the experts diagnosed 43% of these problems to the novices' 15%. On the Single Sentence task, which called for problem representation alone, the pattern is the same. The experts diagnosed 53% to the novices' 23%.

The difference in the amount of expert and novice diagnosis on the 26 "planted problems" is significant even though these include many small and/or local errors we might expect experts, in particular, to handle relatively automatically without explicit diagnosis. When we look at the

amount of diagnosis of more global whole text problems, the contrast between experts and novices becomes dramatic. Figure 10 shows this conttras. In this analysis we looked for an diagnostic statement that showed the writer's concern with a problem that (a) extended over more than one sentence and/or (b) reflected a generalized problem with the text, such as, "this is a sexist letter" or "in a recruiting letter you don't want to focus on what is wrong with the team." Our original text was in fact a hot bed of such global, whole text problems since we had given subjects a text with one purpose, audience, and style (a rather stuffy, disorganized letter to a colleague) and asked them to revise it for a different purpose, audience, genre, and style (a persuasive handout for students). It seemed reasonable to believe students could handle this topic, audience and genre, even if they could not handle revision. Although the novices did draw on their world knowledge and sense of audience and purpose, they appeared to manage the revision process itself by limiting their attempts at diagnosis to the boundaries of a single sentence. As Figure 10 shows, the novices only diagnosed an average of three whole text problems to the experts' ten, each problem receiving a number of comments. (These expert/novice differences were significant at .01 by the Mann-Whitney test.) Even more striking is the amount of processing devoted to these problems once they are noted. That is, once the experts "detected" one of these whole text problems, they followed up with attempts at further diagnosis, the creation of subgoals and text plans, and the use of a variety of other problem-solving strategies. We will illustrate what we mean by a greater "amount of processing" from specific protocols later in this section.

(Insert Figure 10 here)

Whole text problems such as "there's really a contradiction in here" are typically complex, ill-defined problems that may demand both extra attention and some rhetorical knowledge to solve. This study suggests that experts manage the demands of whole text diagnoses by devoting a good deal of attention to the problem; novices manage by avoiding the difficulty altogether.

There are other differences as well in the way experts and novices represent text problems. Let us examine some of the general differences between how our two subject groups represented problems to themselves. In other words, what separates expert and novice diagnosis?

As Figure 9 shows, diagnosis can take at least three importantly different forms depending on whether it is based on a Rule, a Maxim or a sense of the text's Intentions. Intentional diagnoses seem to be the most difficult to perform, because they are based on a comparison of the text and an image of the purpose, audience or point of the text. Unlike Rules or Maxims which are part of the revisor's <u>prior</u> knowledge, these intentions are often <u>created</u> afresh for each text; they may require the writer to make inferences about the audience and to construct a sense of gist or logic of the text. The procedures for solving such problems are also more complex, depending on heuristics rather than rules. In the case of many such problems, there is not even a standard heuristic procedure and the revisor must return to Planning (see Figure 1) and generate new text.

We expected our experts to use a greater proportion of these Intentional diagnoses, assuming they would recognize the hard problems of audience, purpose, paragraph and inter-paragraph coherence and that they would have appropriate rhetorical knowledge to solve these ill-defined problems. Previous research based on analyzing the products of student revision--research in which students' revisions were typically limited to sentence-level corrections with no changes in meaning--would lead one to expect a parallel limitation in the revising process of novices and a failure to consider the more global issues of audience and purpose.

Our results, however, did not support this hypothesis that the process would mirror the product. (See Figure 11). Both experts and novices used Intentional diagnoses with roughly the same frequency. (Given the difference in the overall number of diagnoses, there was no significant difference in frequency: Experts = 69.6%; novices = 63.1%). This finding raised an important question: were the novices actually attending to *global problems* just as the experts were?

(Insert Figure 11 here)

A more detailed analysis of the novices' diagnoses showed that although their efforts were indeed global in their concern for *intentionality*, they were not global in terms of *scope*. Novices regularly demonstrated these inferential responses to the text, to its purpose and audience, but these intentional diagnoses were only applied to problems within the scope of a single sentence, often to single words or phrases. For example, a novice reading the phrase "very fine friendships are often a dividend (of sports) responds: "Wait! This sounds really bad. 'A further dividend'. . . That sounds like some sort of banking instruction or something . . . This 'very' I don't like. . . . This sounds pretty pushy". This novice is clearly attuned to how readers respond to diction and to ill-concealed intentions. However, the chances of making significant changes to the text are probably low if one's unit of analysis is limited to individual words, phrases or sentences.

The ability to deal with global features of the text, it appears then, has two dimensions. One is a concern for *intentionality*--which these novices had in abundance--and the other is the ability to apply that concern to problems of some *scope*. Our initial findings showed that novices largely failed to see problems at the whole-text level (Figure 10). In addition, we found that the overall percentage of multi-sentence intentional diagnoses was 31.6% for the experts versus 10.3% for the novices (See Figure 11).

This interesting surprise in the data suggests that assumptions about the writer's thinking process based on the evidence of the written product may lead to poor inferences about the skills a novice really does possess and about the processes the task really requires. Given this single sentence strategy, urging inexperienced writers to attend to global issues of purpose and audience,

for instance, might do little to affect the part of the process that is really limiting performance. An increased concern with intentionality could be incorporated quite easily into the novice revisor's current, limited-scope strategy with little effect on performance.

The results from the Detection study (Section VI) show that the novices are simply giving themselves far fewer problems to deal with from the very start. As we will now demonstrate, when we look at how the remaining problems are represented (i.e., when we look at bona fide diagnoses on both planted and whole text problems) we see an even further decrease in the amount and quality of information attached to the novice's problem representation. Whereas novices create a spare image of a few problems, experts richly elaborate representations of many. Therefore, given that diagnosis is a prominent part of the expert's process, the next question to ask is, what role does it play?

### What Role Does Diagnosis Play in Expert Revision?

Insofar as literary critics, rhetorical analysts, and writing instructors can serve as role models, one would expect diagnosis to be standard operating procedure for effective revisors. However, the reality appears to be more complex.

1. Diagnosis is clearly not an obligatory step in revision.

Although the Experts successfully fixed 91% of the "planted problems," they chose to actively diagnose only 43% of them. Furthermore, on the diagnostic Single Sentence task itself, they explicitly diagnosed only 53% of these common problems (see Figure 7). Why? The answer is that even experts miss things other experts would see. Secondly, most subjects brought a narrower set of criteria to the Single Sentence task, ignoring many of the more global problems of tone and style they had attended to during their revision. Finally, and perhaps most importantly, some problems appear to be solved more easily by the Detect and Rewrite strategy than by the more demanding path of Diagnose and Revise.

2. Rewriting is often more efficient than trying to Diagnose and Revise.

As Figure 2 shows the simplest path through revision is to simply Detect and Rewrite without calling on the special knowledge or extra attention Diagnosis requires. Experts, like novices, often choose a paraphrase procedure which involves reading a faulty sentence or paragraph, extracting the gist, and then rewriting in the same way they normally produce prose. The Rewrite strategy depends on the writer's normal language production skills. In the case of the Expert these may be considerable. In the case of a novice these skills may be limited, but still considerably in advance of his or her diagnostic skills. Evidence of the rewrite strategy in our study came in many forms: in the speed of revision, the absence of diagnosis, the fluent production of whole new sentences or chunks of text, and in explicit comments by subjects, e.g.,

"I'm rewriting not revising this now" or "What I'm doing is trying to say it in my own words." In the Cued Recall, subjects regularly justified changes with the comment; "I don't know why, this just sounds better." The frequency of the "sounds right" metaphor gave the impression that revisors often work as some musicians play, "by ear." This finding is not surprising since this generate and test process draws on the writers vast knowledge of language patterns and a body of relatively automated production skills.

Diagnosis, then, as we have defined it, appears to be an optional, strategic choice made against the backdrop of the simpler more dominant procedure of Rewriting. These strategic choices will be discussed in detail in the next section of the paper.

3. Diagnosis is the Expert's Option.

Diagnosis is neither obligatory nor as pervasive as the simpler Detect/Rewrite strategy. However, it appears to be a strategy of considerable power that experts have a greater option of using.

To begin with, the Experts are able, when asked, to diagnose more problems. (Compare 53% of the planted problems on the Single Sentence Task to the Novice's 23%). More importantly, experts do significantly more diagnosis, by choice. Figure 11 shows the total number of diagnostic comments made on both the protocols and cued recall. This number includes responses to not only the 26 planted problems, but to any problem the subject perceived. The experts averaged 62 diagnoses over the task to the novice's 30, significant at .05 level. Perhaps an even more meaningful figure emerges when we look at the subjects whose overall strategy predicts that they would use diagnosis if they could. That is, early in the study (and independent of this analysis) we identified four subjects whose explicit definition of the task effectively precluded a diagnostic strategy. As we will discuss later, three subjects (a teacher, professional, and student) worked as rewriters and a fourth (a professional) worked as a deleter. Both of these strategies can increase efficiency by eliminating the need for diagnosis. When we look at the remaining bulk of the subjects, for whom diagnosis is a meaningful option, we find the experts averaging a striking 89 diagnoses over the task to the novices' 33. (See Figure 12)

(Insert Figure 12 here)

One final analysis supports this notion that Diagnosis is an option experts can choose. If we look at those problems which the writers actively detected--ones we can be *sure* they noticed--we find that the experts (including the rewriters and deleter) choose to diagnose 72% of these, while the students only diagnosed 39%.

To sum up, our analysis suggests that although Diagnosis is an optional strategy, it is the Expert's option--and that option is actively exercised. More generally, it points to the fact that revision itself is a strategic process--not a simpler task of detecting and repairing errors. Most of the significant problems in adult texts are ill-defined problems which the writer must represent to himself or herself. In such revision, problem representation is no easy task; it calls on both knowledge and judgment. A Diagnosis, as we have defined it, is a high-information representation, which places the problem in a category, and may call up a repertory of procedures and tests. Diagnosis, it seems, is the expert's option for use on demanding problems. Although it requires more time, attention and knowledge, it is a powerful problem-solving strategy, which, as we might expect, experts choose to use on significant problems.

4. What Constitutes an "Expert" Diagnosis?

Experts not only choose to do more diagnosis, they appear to create qualitatively different problem representations when they do. The following excerpts from the protocols of an expert and two novices will illustrate some of the differences. Notice the various dimensions of the problem noted in the expert's extended representation, compared to novice #1's simpler conceptualization and novice #2's reliance on a rewrite strategy. Notice also that the expert sees problems in relation to the purpose of the whole text, whereas both novices do not look beyond the immediate context of the individual local problems.

> The text, which has been discussing varsity sports, starts a new paragraph with:
>
> I don't want to infer that the only chance women get for participating in sports is on varsity teams. Intra-mural sports are not quite the same as varsity sports in which the rules are better, equipment is better, with the techniques of the players being more developed.

The expert's response is spread over two passes at the text:

```
P8      Aside from the fact that the sentence isn't parallel
P9      Again I don't see the connection between the two sentences.
        [She had seen a connection problem in previous sentences]
P10     It seems to me that if the goal of this article is to
        get women involved...a..
P11     it ought to be clear whether they want them to be involved
        in intra-mural or in varsity sports or both....
P12     right at the beginning.
P13     I'll just go on.
*
*
        (Reads "intra-mural sports" ...intra-mural sports")
P42     I have to go back and look at the purpose of this thing
        once more.
P43     It does not say everything that probably needs to be said
        about student pressure.
P44     The purpose of this thing is to get women into a large
        active group to lobby for more funding,
P45     and the way to do that is to recruit more women for teams.
```

P46 In that case, if we want to recruit more women for the varsity teams, why for intra-mural sports?
P48 I don't know.
P49 I don't know because I don't understand enough about the subject
P50 to know if women in intra-mural sports would be part of what she's calling the teams
P51 and would therefore be part of a group to lobby for more funding
P52a I suppose that they --
P53 if they were interested in sports, even if it were only intra-mural,
P52b they could conceivably still be interested in lobbying for funding.
P54 On the other hand, if the focus here is on varsity teams,
P55 then this paragraph on intra-mural has absolutely no purpose in the thing.
P56 I think I'm going to take the paragraph out
P57 because it really doesn't say anything very helpful about intra-mural sports anyway.
P58 It just says they can be fun and take less time.
P59 It doesn't say anything very positive or convincing for intra-mural sports.
P60 So I'm just going to cross the whole paragraph out.
P61 That takes care of that one.
P62 Now I'll come back to the opening paragraph
P63 and see if I can make it make a little more sense.

Novice #1

(Reads "I don't want...")
P41 This is all negative stuff
(Reads "in which the rules are better, equipment is better, etc")
P42 It can all be cut out
P43 (Writes "I.M. sports are not the same as varsity sports".) (Speaks) Period.
P44 They (are) a little casual, I guess. (She later adds this idea to her rewrite of the sentence)

Novice #2

P103 O.K., so--a--that's bad
P104 Looks like a total rewrite

(Reads "I don't want to infer.."
P111 That's bad

Novice #2 chose to avoid diagnosis and take a direct path from Detect to Rewrite. Novice #1 chose to diagnose the problem she detected in the first sentence and to revise both the sentence and the information it contained. The expert chose to exercise her option to diagnose, building a quite complex representation of the problems with this paragraph.

We believe that further research will show other qualitative differences in expert diagnoses, not merely differences in frequency. As the expert protocol excerpt above shows, experienced writers use diagnosis to explore ill-defined problems--problems for which they, like the novices, have no ready-made procedures or heuristics. This expert created an elaborate problem representation which included information about parallelism, about logic, topic focus and coherence, as well as her inferences about the alternative purposes of the text and her plans for meeting those purposes. Her diagnosis, however, is not a mere list of faults, but a hierarchically structured representation of the problem. For instance, she interprets the local problem of a topic shift between sentences 1 and 2 (P42) as a failure to maintain the purpose and focus of the entire handout. Her diagnosis of local incoherence between sentences is subordinated to the problem of global coherence. Likewise her solution is based on converging evidence--the text is incoherent and it is low on information. Note too how her abstract diagnoses concerning purpose are clearly connected to specific features of the text. This ability to build elaborated representations that integrate diagnoses into networks of related problems at various levels of the text appears to be an important feature of expert performance. (Flower, Carey and Hayes, in preparation). Diagnosis in the hands of an expert leads to the creation of a substantially different problem from that which the novice sees.

Diagnosis appears to be the expert's option for good reason. It is an exceptionally powerful means for managing complex problems. However, that power often comes at a price. On the one hand, diagnosis allows writers to grapple with larger more global problems; it consolidates problematic features of a text into a conceptual whole, and it points to ways of dealing with the whole text. On the other hand, as writers begin to detect and represent these more global problems to themselves, they move from local and more well-defined problems with explicit procedures to more complex, ill-defined problems for which there are no clearly defined ways to proceed. Diagnosis is a double-edged sword. Representing a complex and global problem is often tantamount to giving yourself an ill-defined and unmanageable problem. In further studies of diagnosis, we hope to see how writers handle this dilemma and use diagnosis to create manageable problems.

# VIII. STRATEGY SELECTION

Revising prose requires the skill to construct solutions to problems--solutions that in their particulars are unique for a given task. Because many text problems are ill-defined, writers usually have a range of available means for solving them. Writers vary widely, however, in the problems they choose to solve, in the strategies they use to solve problems, and in the quality of solutions they create.

Novices, as one may expect, have an impoverished set of problem solving strategies, tending to use a few rule-governed procedures to solve most problems they detect. Aside from invoking rigid rules (Rose, 1980), novices deal with many text problems by using a default strategy of simply "eliminating the obstacle by deleting it" (Sommers, 1980; Bridwell, 1980). Experts, on the other hand, use a range of solution procedures that allow them to satisfy a more complex set of goals and constraints than novices.

Research into the revision processes of experienced and inexperienced writers offers converging evidence that revision requires strategic decisions about how to modify text. Early research by Beach (1976) found that "extensive revisors could predict changes necessary for subsequent drafts. Having described a writing problem, they would predict ways of solving that problem on their next draft." Non-extensive revisors were less likely to predict solution strategies or to make many changes on later drafts. Beach's findings implied that if students can be taught to revise more extensively, their drafts will improve. But Bridwell's (1980) findings contradicts Beach's assertion that more extensive revisors are better writers. Her study points out that more revision doesn't necessarily mean better revision: "the most extensively revised papers received a range of quality ratings from the top to the bottom of the scale." In fact, her results show that novice revision can often make the text worse. Similarly, Bracewell (1983) found that novices often lack the ability to exercise deliberate control over the syntactic productions they make--that there is often a disassociation between goals and actual performance. Success in revision, then, appears to lie, not in the amount of revising one does, but in how revision is accomplished.

We find that the best revisions are produced by writers whose strategies are adapted to the texts' purpose and goals. But what does it mean to make this adaptation? In this section, we will discuss a range of strategies writers may have available to them during revision--strategies which help give writers control over their revision process and strategies for modifying text. We will also see how strategy selection reflects changes in the writer's task definition.

**STRATEGY SELECTION** encompasses those critical decisions writers must make about what action to take after they have defined the task, evaluated the text or plan, and represented its problems. This decision process is represented by the **select strategy** component of the model. Of course, decisions are made throughout the revision process. Our focus here is on the writers' active strategies for monitoring the writing process and for modifying text. When writers define their task, evaluate the text, and represent problems in light of goals and plans, there are at least five actions from which to choose--choices that help the writer to manage the revising process:

- **ignore** (the problem)
- **delay** (effort to solve the problem)
- **search** (for more information to solve the problem)

and strategies for modifying text:

- **rewrite** (the text by redrafting or paraphrasing)
- **revise** (the text by relying on diagnostic information).

The first three actions often assist the writer in managing the process, in clarifying the problem representation, and in determining the scope of particular text problems. The second two describe major paths that lead to modifying text: the rewrite or revise strategy. Our results suggest that expert and novice revisors are quite distinct in their ability to exercise these strategies.

## Managing the Revising Process: To Ignore, To Delay, To Search

### Ignore

The first action on the left side of the **select strategy** area, **ignore**, describes what writers do when they detect text problems that they believe are not serious enough to worry about. Typically, experienced writers ignore detections associated with optional grammatical or stylistic rules such as "never end a sentence with a preposition." Maxine Hairston's (1981) study on style points out that often grammatical and stylistic rules are applied arbitrarily and frequently are influenced by personal taste.

Our results support Hairston's findings; not only is there a wide difference between experts and novices in the range of text problems writers pay attention to and ignore, but expert differs from expert and novice from novice. For example, one expert in our study noticed a split infinitive and said: "I see we have a split infinitive here, but they never bother me so I think we'll leave it." Another expert who spotted the same split infinitive remarked, "I always feel compelled to fix split infinitives, I just hate the way they sound." On the other hand, a novice who noticed a problem with word choice said "I know its not quite the right word, but I won't

change it; they'll never notice it anyway, so why bother. I'll just go on to the next sentence." And another novice who detected a problem with an undefined abbreviation read it and commented "I wonder what the term 'AIAW' means? I can't figure it out, but it must be important, so I'll leave it there."

Moreover, we found that expert and novice writers often distinguish the text problems they ignore using one of two criteria: 1) the problem will not create confusion for the reader; or 2) finding a solution is too difficult or not worth the effort. Sometimes both criteria are used simultaneously. Making effective decisions about what to ignore depends on skill in knowing when to apply these criteria--skill that requires making what are often subtle discriminations. And the ability to make these discriminations is one factor that separates novices from experts. We find that *ignoring* can be a sophisticated strategy for controlling the revision process, but that too often it acts as a default strategy in the hands of inexperienced writers.

Delay

To **Delay** action describes the writer's conscious decision to divide the revision process into parts. Delaying action allows the writer to focus attention selectively on one part of the task while *setting a goal* to deal with another part later. The decision to set the goal of "I'll deal with this text problem later" occurs when:

- The writer chooses to prioritize revision activity--dealing with a class of problems such as those associated with the text's meaning first, with grammatical and stylistic problems second, and with polishing last. Experts, as we pointed out in our task definition study, often read a text for gist before repairing surface structure features.

- The writer may need to reread the text to evaluate whether the problem is major or minor in relation to other modifications that need to be made. We saw that expert writers frequently re-evaluate the problems detected during a first reading to determine the order in which problems will be addressed.

- The writer may not have a ready procedure for remedying the text problem and chooses to "write around" the difficulty until the solution becomes apparent. One of the better novices in our study noticed that the original text didn't say much about sports being fun and couldn't decide how to convey that idea. She made a note to herself "include why sports are fun"--"as soon as I figure out where to put it and how to say it."

- The writer realizes that invention is called for, but needs to evaluate the text more critically in order to discover the best way to proceed. For example, several of the experts in our study decided that a revised handout "on sports for freshman women" should list the sports the college offers. To determine what kinds of sports they would include, they put off generating the list of sports until they had decided on the argument's main points.

- The writer needs to revise a particular section of text before knowing how to solve problems in other parts of the text. For example, writers frequently revise sections

in the body of a text before revising introductions. One expert in our study remarked "if this is a handout, we need to get the facts straight first, then I'll know what to do with this deadly introduction."

As we can see from these examples, to delay action often sends writers back to planning, re-evaluating, and reformulating goals for the text. Both expert and novice writers tend to create verbal cues for themselves, such as "develop this point as soon as I see how it relates to the next section," as well as visual cues such as checking or starring sections of text as reminders of changes they need time to consider.

To delay action effectively requires the writer to attend to text features that need further work. We found that while revisors often intend to come back to places in the text that require more attention, they frequently fail to execute these intentions. For example, one of the novices in our study noticed that the original text set up a a negative contrast between intramural and varsity sports, implying the superiority of varsity sports. She remarked "if this handout is to recruit people for both kinds of teams, maybe it shouldn't say that varsity sports have better players, rules and equipment. That's like saying, you know, your equipment is going to be worse, and the rules are going to be worse, and you may get hurt." She then said, "well, I'll come back to that and change it later. I think it should make both kinds of teams sound good." Her final revision, however, kept the pejorative contrast.

While the reason revisors sometimes fail to act on problems they flag is not entirely clear, we offer several hypotheses:

- they may lack the strategies and procedures for acting on their goals
- they may lack the managerial skill to coordinate a number of revisions
- they may lack the perspective to judge the appropriateness of a proposed solution
- they sometimes experience memory overload and forget the problems they detect
- they sometimes ignore their early detections and satisfice with revisions already-made
- they may feel that the effort needed to make the revision is too great for the impact it will have on the text.

Search

**Search** is a strategy used to refine and elaborate the problem representation. Writers typically search as a response to a feeling that their representation is incomplete or inadequate. Search is a process of re-representation in which writers probe to define the problem more explicitly. The output of problem representation, as we have suggested previously, is a continuum from ill to well-defined problems. In using a search strategy, writers try to push their

representation from ill to well-defined. This process provides writers with a sharpened representation, which in turn gives them a better idea of how to solve a problem.

To re-represent a problem involves many processes; prominent among these are: *memory search* and *text search*. By *memory search*, we are referring to the process of retrieving relevant experience and knowledge associated with the problem. Memory search often occurs rapidly. Writers may detect a problem and characterize it generally, but then generate several successive re-representations of its nature. Consider the following example of a novice who searches memory to sharpen his problem representation, finds relevant and useful information, yet fails to accomplish his revision goal.

The writer has just read a sentence describing a lack of competition in varsity sports: "Therefore, many of the teams are less competitive than many students expect." He responds, " Wait a minute, there's something wrong there. I don't know if I want to say that sports are less competitive. Is this true anyway? Well, when I played varsity sports, I noticed that students were *more* competitive than I'd ever seen before. So that's wrong. Maybe it should say less strenuous. . . no. . . maybe less rigorous. You know, some women may be a little intimidated to come out. How about 'teams are not less competitive.' No, that sounds bad. I don't know how to get across the fact that they're less than major college sports but they're just as competitive. O.k., I'll just cut that whole thing out."

In this case, the subject found that information was imprecise, but was unable to increase its precision as he wanted. We find that although experts are typically successful in using information retrieved from memory to improve the text, novices tend to fail to use retrieved information effectively.

*Searching text*, on the other hand, is a process in which the revisor scans the text to get a better understanding of a text problem. We observed that experienced writers sharpen their representation of a text problem by rereading with a focused goal in mind. Rereading with a focused goal helps the writer both to further specify the problem and find all related instances of that problem. Writers engaging in this re-representation of the text create goals that respond to both global and local features of the text. Some of the goals we saw writers create prior to searching the text include:

- pinpoint the location of contradictory assertions
- check for missing or redundant information
- evaluate for organization and coherence
- locate arguments that may offend the audience

- check for the consistency of subheadings and other internal structural markers
- check for variety of sentence openers and sentence structures
- check for the frequency a particular phrase, word, or concept has been used

We find that writers often employ a text search strategy to determine the location, frequency, and scope of problems to be found in the text. However, many times writers search not for errors of commission, but rather for what is missing. They look consciously for gaps, missing information, and places where examples, analogies, or metaphors are needed.

For many experienced writers, search is a powerful strategy for discovery and subsequent invention. As indicated by our model of Reading to Evaluate, (see figure 5) there is a wide range of possible outputs associated with the various levels of the comprehension process. Words often suggest alternative words, and schemas alternative schemas. In evaluating text the reader is stimulated to think of new and better ways of saying the same ideas.

Moreover, many re-representations involve the coordination of memory and text search. For example, revisors are sometimes alerted to contradictory assertions when something currently being read doesn't fit with what they remember of preceding text. In this situation the writer initiates a search of the text for that earlier information.

## Strategies For Modifying Text: Rewriting and Revising

We observed two distinct strategies that lead to modifying the text: **REWRITING** and **REVISING**. Our characterizations of rewriting and revising strategies grow out of analyzing how writers carry out the task they define for themselves. For the most part, our writers chose one of these revision "paths" and stuck with it. By "path" we mean a consistent set of revising strategies, whether by design or default, that describes the majority of the writer's responses to text problems. As one can imagine, elements of both strategies appeared in varying degrees in the protocols of both expert and novice writers.

### Rewriting: Redrafting or Paraphrasing

**Rewrite** is a strategy used by revisors who are not necessarily trying to save the original text, but only the meaning. Rewriting generally takes one of two shapes: **redrafting** or **paraphrasing**.

When revisors **redraft** a text, they often extract the text's gist and then abandon the existing text and start over--sometimes rethinking the top level goals and modifying text plans; e.g., what I intended was to inform my audience by making three points, but I want the audience to see how these points are shared values, so I need to argue these points by creating more of a

shared context rather than to simply list my points. The writer then recasts the entire argument structure. The redraft strategy is thus a restructuring of the text's purpose or plan. It does rely on transformations of sentences or clauses to produce new text.

People redraft, for example, when the existing text is riddled with both high and low level problems, inconsistencies that compete for the writer's attention, and so the revisor saves effort through *not diagnosing* each problem, but rather by making a whole text diagnosis such as "this text will not accomplish my goal for the reader, I need to rewrite the whole thing."

**Paraphrase**, the second rewrite strategy we observed, is invoked when the writers goal is to save the existing text's meaning, but to express the meaning in different words, using a different surface structure. Novices are more likely to paraphrase than to redraft, partly because the implicit test of paraphrase is a comparison between it and the integrity of the gist and not a check against a plan.

Writers usually paraphrase the text by extracting the sentence by sentence gist and then recasting the sentences. The writer who chooses to paraphrase and rewrite often does so in a linear manner. This strategy is accompanied by considerable amounts of rereading the text produced up to the last change. Like redraft, paraphrase uses little if any specific diagnostic information, but relies mainly on making detections and in seeing the gist.

We observed that experts often use a "density of text problems" heuristic to prompt a redraft revision strategy. By "density of text problems," we mean the frequency and diversity of text problems. Using a density heuristic tells the writer to attend to the frequency and type of text problems before taking action. This decision procedure is useful because it allows the expert to: 1) coordinate the solution of separate text problems to produce a coherent revision; and 2) to select efficient revision strategies. Further, it provides the writer with criteria for deciding whether to accept the text's surface structure or to reject it and start over. We have found that experts using such an heuristic have the advantage of making decisions that are better informed by the requirements of the task. For example, when one of our professional writers noticed that the opening paragraph of the letter was the wrong tone for the intended audience, her reaction was to scan the text to see how many times a word or phrase offended the audience--she then used the density guideline of, "I can't simply change a word here and there, because the whole tone is offensive, I'll have to rewrite the whole thing."

Another situation that prompts a rewrite strategy is when writers run into a whole range of problems that would take considerable effort to fix one-at-a-time with the help of diagnoses. Experienced writers in this situation may say to themselves, "I see too many different problems in here and I'll simply need to rewrite." There are many reasons why people choose to rewrite rather than repair the existing text. Below are a few:

- when writing for discovery, rewriting can free the writer of the existing text's influence on production of new ideas or ways of presenting ideas
- when the writer doesn't have an adequate strategy for fixing the text problem
- when the writer is unable to integrate the text's gist or when there are several competing gists that may cause the reader difficulty

While rewrite can be a powerful strategy when the situation calls for it, we have observed problems such as these associated with rewriting:

- it can be an inefficient strategy for novices because they frequently do not have sufficient criteria to decide when to accept their rewrite and therefore don't know when to stop generating new versions. Because it requires so much trial and error, it usually takes longer than a focused a solution.
- sometimes rewriting can be a very frustrating activity. When the writer is not especially fluent, even the rewriting of a sentence can be cumbersome. When novice writer gets bogged down with a particular sentence part, it becomes difficult for them to see the context in which the sentence functions.
- sometimes the existing text is a powerful misleading stimulus, one that short circuits invention and leaves the writer with a myopic perspective of the text. When the novice writers tries to paraphrase the existing text, the recasting of the old meaning interferes with producing the intended one. The novices attention to the existing text limits the alternatives considered. In trying to recast the same poorly formed syntactic and semantic aspects of a text, it becomes difficult then to decide whether the new sentence fits the context better than the old.

We found a special class of writers whose choice of goals distinguish them from the rewriters and the revisors: the **deleters.** Deleters are writers who systematically excise text problems rather than invent, restructure, and use procedures for remedying text difficulties. Unlike the rewriters, deleters change the surface structure of the text by eliminating problematic words, phrases or sentences, and in so doing, they often lose much of the propositional content of the text. Deletion is the quickest way to modify text, but if driven without clear goals, deletion can make the text worse. When the text needs little improvement, however, deletion is an efficient strategy.

Revising

The second major path we saw writers choose is that activity of making diagnoses and then selecting a fix procedure--a path we call **Revise.** Our model presents "revise" in a special sense: "revise" refers to the process of deciding that you will use diagnostic information to fix text problems. Writers usually revise when they feel that much of the text can be saved. It does not exclude, however, that revisors often diagnose gaps in text that call for elaboration, examples, details.

We can see how experts and novices differ in the problems they perceive and in the procedures they use for improving text by looking at their problems and solutions paths diagrammatically via a **Means-Ends Table.** We are suggesting that the output of revising is a matrix of problems and solutions.

## Means-Ends Tables of Revisors

Means-ends analysis was described by psychologists Newell and Simon (1972) as a general problem solving procedure in which symptoms of problems were matched to strategies for solving them. For any writer, we can construct a means-ends description of the problems represented and the strategies taken to solve the problems. Means-ends tables for individual revisors then are not theories, but descriptions of the problems the writer defines and the actions the writer uses to correct them.

Figures 13 and 14 show two sample means-ends tables, one of a novice and the other an expert. In the left column are the ends, or the problems the writer flagged. (Novice writers did not always use these particular diagnostic categories in naming the text problems; they were inferred by two separate coders.) Along the top of the table are the means or the strategies the writer used to solve the problems on the left. We can see that for this novice, the detection of global problems is quite limited. More importantly, novices seem to adhere to the same strategies for remedying a range of text problems with deleting, improving wording, and constructing new material the most frequent means for solving text problems. Further, notice that the novice is typically working with text units no larger than two sentences.

Insert Figures 13 and 14 here

Compare this table to that of an expert: notice the elaboration of the local and global ends. We see a broader range of global problems attended to, including gaps in text content that call for inventing paragraphs, examples, and transitions. The experts' means for solving these difficulties is also more elaborated and more concentrated on paragraphs than the novice. The expert is also a better manager of the revising process. Notice how this expert chooses to "delay action" and to "search for information" as a problem-solving activity for global difficulties.

We observed experts using what we call a "precedence guideline" of handling more important problems before spending time on minor ones. When rereading their text to see if, for example, transitions are logical, and they find a disorganized paragraph, they will stop looking at the transitional phrases and concentrate on revising the disorganized paragraph. This inventory-making skill places the expert writer in a position to decide whether to rewrite or revise the text--a useful advantage.

When we look at the "local" part of the experts' means-ends table, we again see that it's more developed and contains embedded categories. The expert notices specific subsets of problems within larger categories, distinguishing between ambiguity and redundancy, for example. The expert also has a wider range of procedures for dealing with local problems. Notice how specific the procedures get--use infinitives, insert qualifiers, choose synonyms.

Writers who exhibit a single means for solving text problems--we've called them deleters--have the sparest of all possible means-ends tables. They use "cut and slash" as their means for solving all text problems. As one might expect, the extremely limited problem-solving means used by deleters doesn't always produce the best revisions.

### Major Revision Paths

Although the rewrite and revise paths are handled quite differently by experts and novices, we can't argue that one path is generally superior to the other. We found that expert performance is marked by the ability to make strategic decisions about which path to choose given the rhetorical situation, the quality of the text produced so far, and the pragmatic constraints under which the revision activity takes place. We observed that novices tend to choose one path and stick with it, while experts sometimes opt to switch paths when needed, e.g., two of our experts took a diagnosis-revise path first and then decided the text was too flawed to fix and rewrote it.

We found that writers often choose a revision path on economic grounds, that is, they choose the path which appears to allow them to meet their goals with the least effort. The perceived effort required by **revising** depends on at least two factors:

- The perceived quality of the text, that is, the kind, number, and scope of its problems.
- The writers perception of his/her competence to fix text problems.

The perceived effort required by **rewriting**, on the other, hand depends mainly on the writer's sense of fluency in generating new text. While novices have difficulty with either revising or rewriting, they have relatively less difficulty with rewriting, and hence often choose the rewriting path. Experts, on ther other hand, because they have facility in both rewriting and revising, can switch readily from one to the other.

We noted that some of our expert writers characteristically rewrote the text, while others always revised. When we evaluated the final products, expert rewriters and revisors did equally well, although they produced quite different kinds of texts (inter-rater reliability 81%). Expert writers then, whether rewriters or revisors, recognize available options and invoke strategies that are best suited for the problems encountered in their current situation.

Overall, our research suggests that we may have the most impact on the revising process of inexperienced writers if we find ways to give them a bigger repertory of options--options that will give them flexibility and increase their problem-solving power.

## References

Bartlett, Elsa J. (1982). Learning To Revise: Some Component Processes. In M. Nystrand, (Ed.), What Writers Know: The Language, Process And Structure Of Written Discourse New York: Academic Press.

Beach, R. (1976). Self-evaluation strategies of extensive revisers and nonrevisers. College Composition and Communication, 27, 160-164.

Bereiter, Carl and Scardamalia, Marlene (1981). From Conversation To Composition: The Role Of Instruction In A Developmental Process. In R. Glaser, (Ed.) Advances In Instructional Psychology, Volume II Hillsdale, NJ: Lawrence Erlbaum.

Bracewell, Robert J. (1983). Investigating The Control Of Writing Skills. In P. Mosenthal, L. Tamor and S. Walmsley, (Eds.) Research On Writing: Principles And Methods New York: Longman's.

Bridwell, L. S. (1980). Revising strategies in twelfth grade students: transactional writing. Research in The Teaching of English, 14 [3], 107-122.

Cooper, Marilyn & Holzman, Michael (1985). Counterstatement Reply. College Composition and Communication, 36, 97-100.

Emig, Janet (1971). The Composing Process Of Twelfth Graders. Urbana, Illinois: National Council of Teachers of English.

Faigley, L., & Witte, S. (1981). Analyzing revision. College Composition and Communication, 32, 400-414.

Flavell, John H. (1977). Cognitive Development. Englewood Cliffs, NJ: Prentice-Hall.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. College Composition and Communication, 32, 365-387.

Flower, L., & Hayes, J. R. (1981). The pregnant pause: an inquiry into the nature of planning. Research in The Teaching of English, 15 [3], 229-243.

Flower, L., & Hayes, J. R. (1981). Plans and the cognitive process of composing. In C. Frederiksen, M. Whiteman, & J. Dominic (Eds.), Writing and the nature, development and teaching of written communication, Hillsdale, N. J.: Erlbaum.

Flower, L., Carey, L., Hayes, J. R. (in preparation). Diagnosis in Revision: The Expert's Option.

Flower, L., Hayes, J. R., Carey, L., Schriver, K., and Stratman, J. (forthcoming). Detection, Diagnosis, and the Strategies of Revision. College Composition and Communication.

Hairston, Maxine. (1981). Not All Errors Are Created Equal: Nonacademic Readers in the Professions Respond to Lapses in Usage. College English, 43 [8], 794-806.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. Gregg and E. Steinberg (Eds.), Cognitive Processes in Writing: An Interdisciplinary Approach. Hillsdale, N.J. Lawrence Erlbaum.

Hayes, J. R., Schriver, K., Flower, L. (in preparation) Decisions in Revision.

Hayes, J. R., Schriver, K., Flower, L. (in preparation) Perceptual Processes in Revision: Seeing Problems in Text.

Murray, D. M. (1978). Internal revision: a process of discovery. In Charles R. Cooper and Lee Odell (eds.), Reseach on Composing: Points of Departure, Urbana, Illinois: NCTE.

Newell, A., and Simon, H. A. (1972). Human Problem Solving. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Nold, Ellen. (1982). Revising. In C. Frederiksen, M. Whiteman & J. Dominic, (Eds.) Writing: The Nature, Development And Teaching Of Written Communication Hillsdale, NJ: L. Erlbaum.

Pianko, Sharon. (1979). Description of the Composing Process of College Freshman Writers. Research in the Teaching of English, 13, 5-22.

Rose, M. (1980). Rigid rules, inflexible plans, and the stifling of language: a cognitivist analysis of writers block. College Composition and Communication, 31 [4], 389-401.

Scardamalia, M. and Bereiter, C. (1983). The development of evaluative, diagnostic and remedial capabilites in children's composing. In M. Martlew (Ed.), The Psychology of Written Language: A Developmental Approach, London: John Wiley and Sons.

Schank, R. and Ableson, R. (1977). Scripts, Plans, Goals and Understanding. Hillsdale, NJ: Erlbaum.

Schriver, K. (1983). Revising Computer Documentation for Comprehension: Ten Exercises in Protocol-Aided Revision. Digital Equipment Corporation, in preparation.

Simon, Herbert A. and Ericsson, K. Anders (1984). Protocol Analysis: Verbal Reports As Data. Cambridge, MA: M.I.T. Press, Bradford Books.

Sommers, Nancy. (1980). Revision strategies of student writers and experienced writers. College Composition and Communication, 31, 378-387.

Stallard, Charles. (1974). An Analysis of the Writing Behavior of Good Student Writers. Research in the Teaching of English, 8, 206-218.

Suppes, Patrick. (1974). The Place Of Theory In Educational Research. The Educational Researcher 3, No. 6, 3.

Thibadeau, R., Just, M., Carpenter, P. (1982). A Model of the Time Course and Content of Reading. Cognitive Science, 6, 157-203.

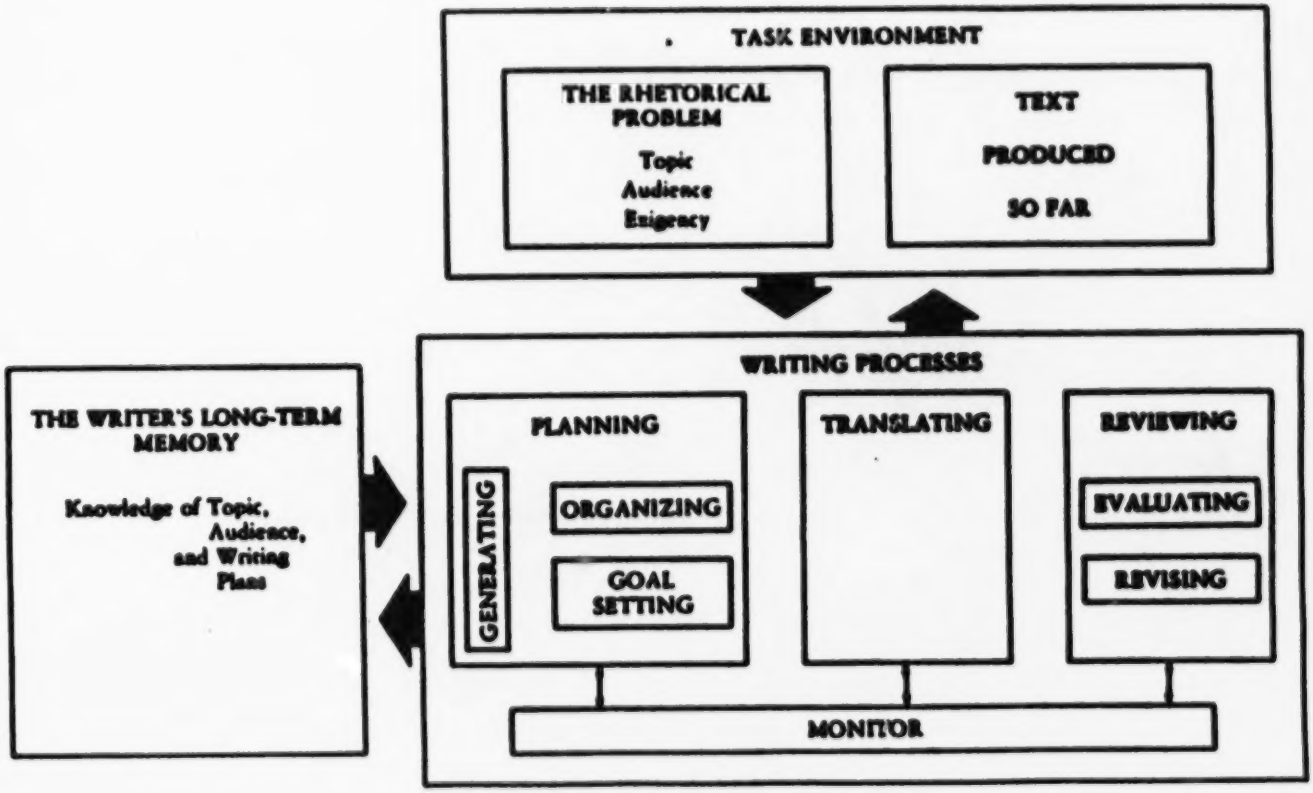


Figure 1. The Structure of the Writing Model.

## COGNITIVE PROCESSES IN REVISION

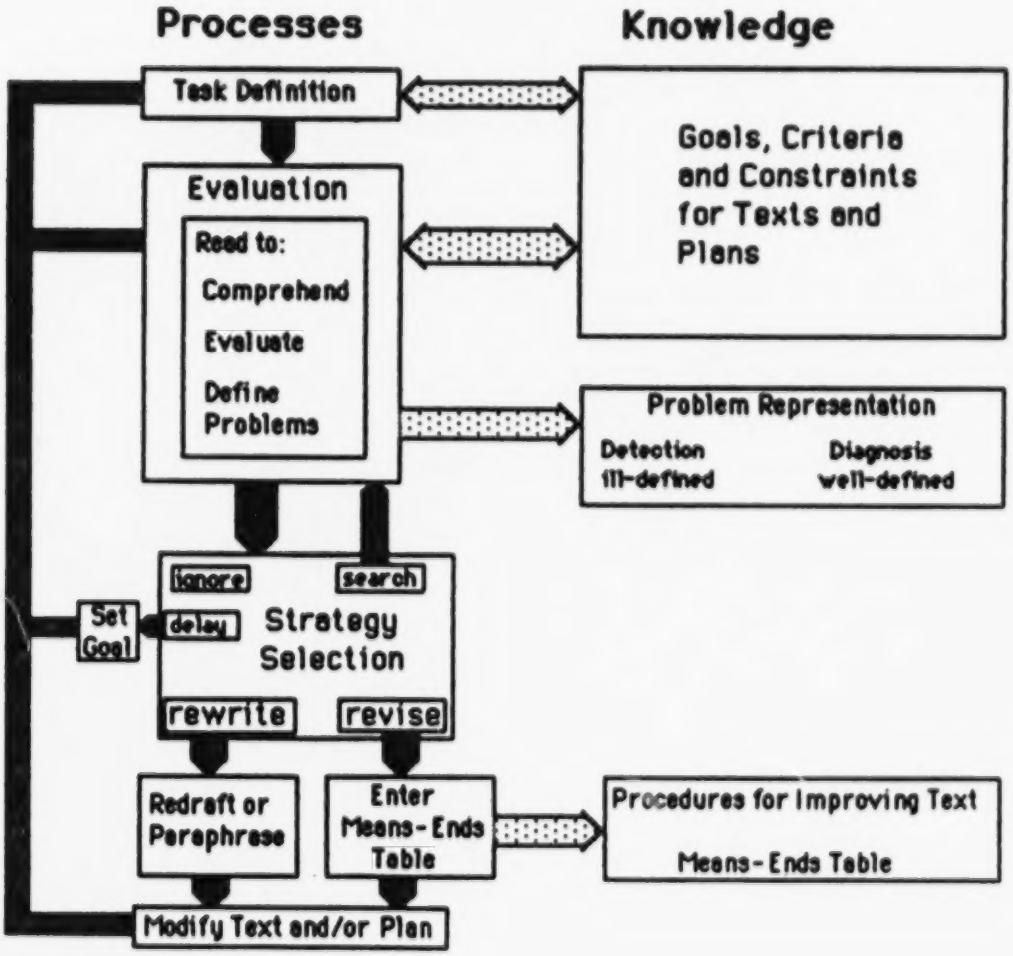


Figure 2. A Process Model of Revision.

| TASK DEFINITION ACTIVITY | EXP | NOV | | NG | GB | KC | ML | MM | CC | JJ | | HD | DE | HL | DS | SE | BA | JP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| READS LETTER THRU BEFORE ATTEMPTING TO REVISE | 6/7 | 5/7 | | X | X | 0 | X | 0 | X | X | | X | 0 | X | X | X | X | X |
| ESTABLISHES PROCESS PLANS OR GOALS THAT ARE MORE SPECIFIC THAN DIRECTIONS BEFORE REVISING | 4 | 1 | | 0 | 0 | 0 | 0 | 0 | X | 0 | | X | 0 | X | 0 | X | 0 | X |
| OFFERS "GIST" OF LETTER CONTENT OR PURPOSE | 2 | 0 | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | X | 0 | X | 0 |
| ESTABLISHES "AUDIENCE" NEEDS BEFORE ATTACKING 1ST SENTENCE | 2 | 1 | | 0 | 0 | 0 | 0 | 0 | X | 0 | | 0 | 0 | 0 | 0 | X | X | 0 |
| OFFERS INVENTORY OF PROBLEMS DETECTED IN FIRST READ | 2 | 0 | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | X | 0 | 0 | 0 | 0 | 0 | X |
| DETECTS "GLOBAL" PROBLEMS DURING OR JUST AFTER 1ST READ (OTHER THAN "AUDIENCE" PROBLEM) | 5 | 2 | | 0 | 0 | 0 | X | 0 | X | 0 | | X | 0 | X | X | 0 | X | X |
| MAKES CRITICAL COMMENTS (OTHER THAN "GLOBAL" OR "AUDIENCE" DURING/JUST AFTER 1ST READ | 5 | 2 | | 0 | 0 | 0 | X | 0 | X | 0 | | 0 | 0 | X | X | X | X | X |
| TOTAL NO. OF INSTANCES | 20 | 6 | Σ | 0 | 0 | 0 | 2 | 0 | 4 | 0 | | 3 | 0 | 3 | 3 | 3 | 4 | 4 |
| | | | | (NOVICES) | | | | | | | | (EXPERTS) | | | | | | |
| PERCENTAGE OF TOTAL POSSIBLE | 47.6 | 14.4 | | | | | | | | | | | | | | | | |

Figure 3. Task Definition Activities in Experts and Novices.

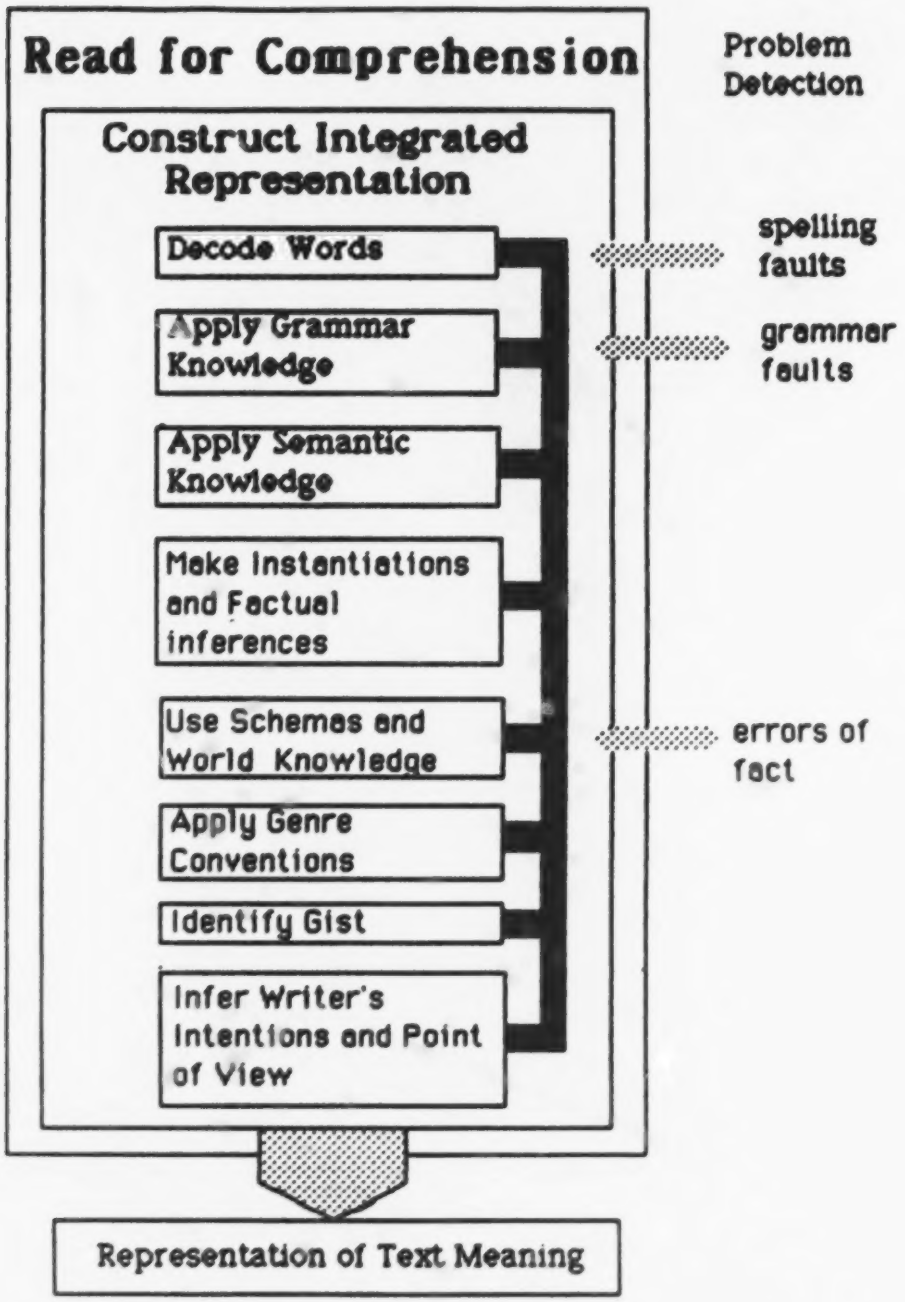


Figure 4. Reading for Comprehension.

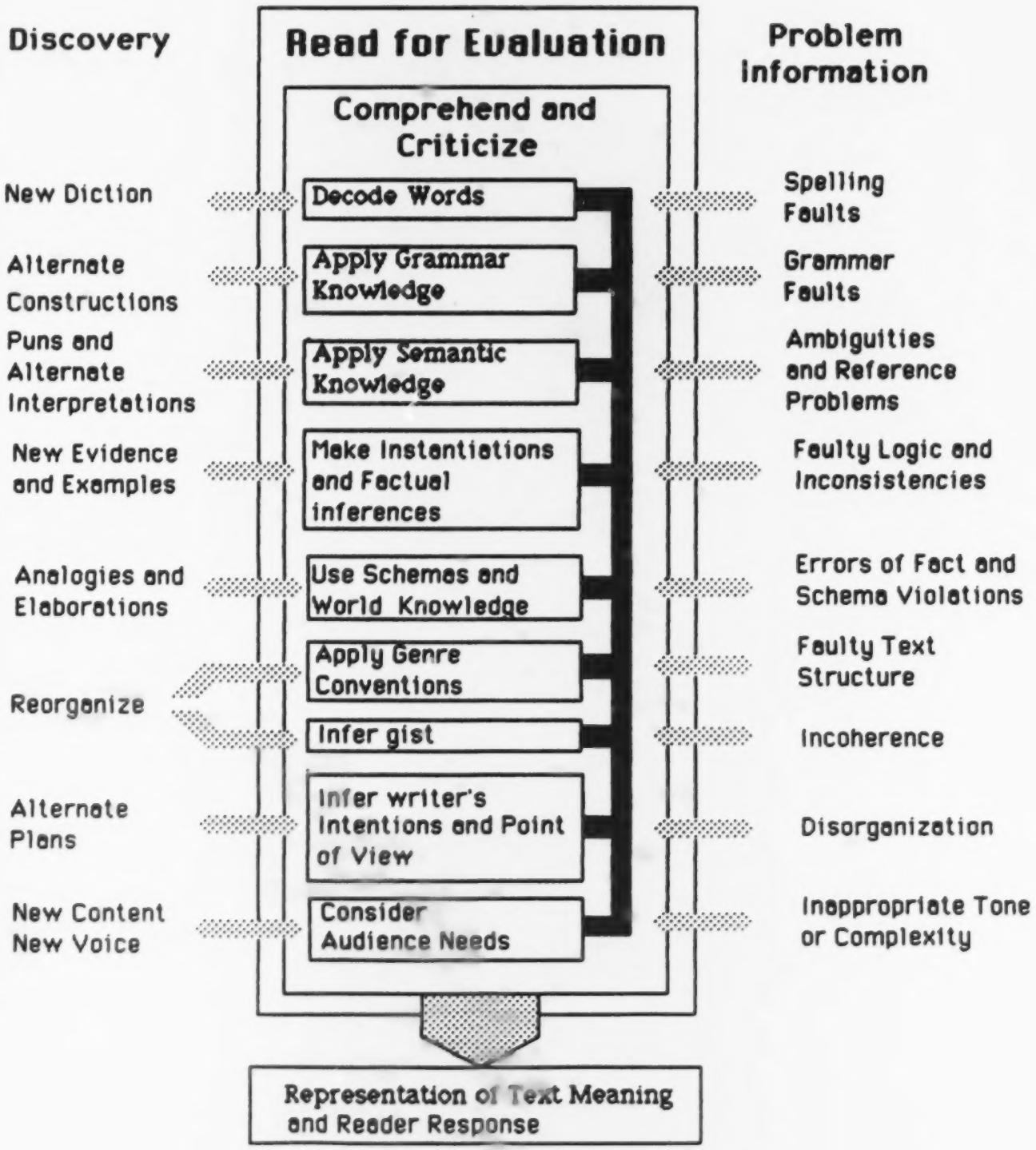


Figure 5. Reading to Evaluate and/or Fix Text Problems.

| | Read | Catch | Approve | Miss |
|---|---|---|---|---|
| Frequency | 327 | 85 | 20 | 222 |
| | | | Σ = 242 | |
| Percent | 100% | 26% | 6% | 68% |
| | | | Σ = 74% | |

Figure 6. Efficiency of Problem Detection in Novices.

| Subjects | Protocol or Cued Recall | | Single Sentence Task | | Fix/Eliminate |
|---|---|---|---|---|---|
| | Detection | Diagnosis | Detection | Diagnosis | |
| **Teachers** | | | | | |
| JP | 18 | 14 | 21 | 20 | 26 |
| SE | 9 | 9 | 18 | 13 | 25 |
| BA | 20 | 20 | 17 | 16 | 23 |
| DS | 18 | 11 | 20 | 10 | 24 |
| **Professionals** | | | | | |
| DE | 9 | 2 | 15 | 12 | 25 |
| HL | 14 | 8 | 9 | 7 | 23 |
| HD | 18 | 14 | 21 | 19 | 20 |
| Σ Experts | 106 | 78 | 121 | 97 | 166 |
| % possible | 58* | 43** | 66 | 53 | 91 |
| **Experienced Students** | | | | | |
| CC | 8 | 5 | 11 | 5 | 26 |
| JJ | 12 | 4 | 15 | 12 | 14 |
| **Novice Students** | | | | | |
| NG | 13 | 6 | 10 | 5 | 13 |
| GB | 3 | 1 | 9 | 2 | 12 |
| MM | 6 | 2 | 8 | 4 | 21 |
| ML | 12 | 6 | 13 | 8 | 16 |
| KC | 12 | 3 | 11 | 6 | 15 |
| Σ Students | 66 | 27 | 77 | 42 | 117 |
| % possible | 36* | 15** | 42 | 23 | 64 |

* Expert-novice difference significant at .02 level.
** Expert-novice difference significant at .01 level.

Figure 7. Detecting, Diagnosing, and repairing text problems.

| | | Local | Global | Other | Total |
|---|---|---|---|---|---|
| Expert | Number | 54 | 42 | 5 | 101 |
| | Percent | 54 | 42 | 5 | |
| Novice | Number | 71 | 13 | 3 | 87 |
| | Percent* | 82 | 15 | 3 | |

Figure 8. Episodes Dealing with Local and Global Problems in Experts and Novices.

## Problem Representation Continuum

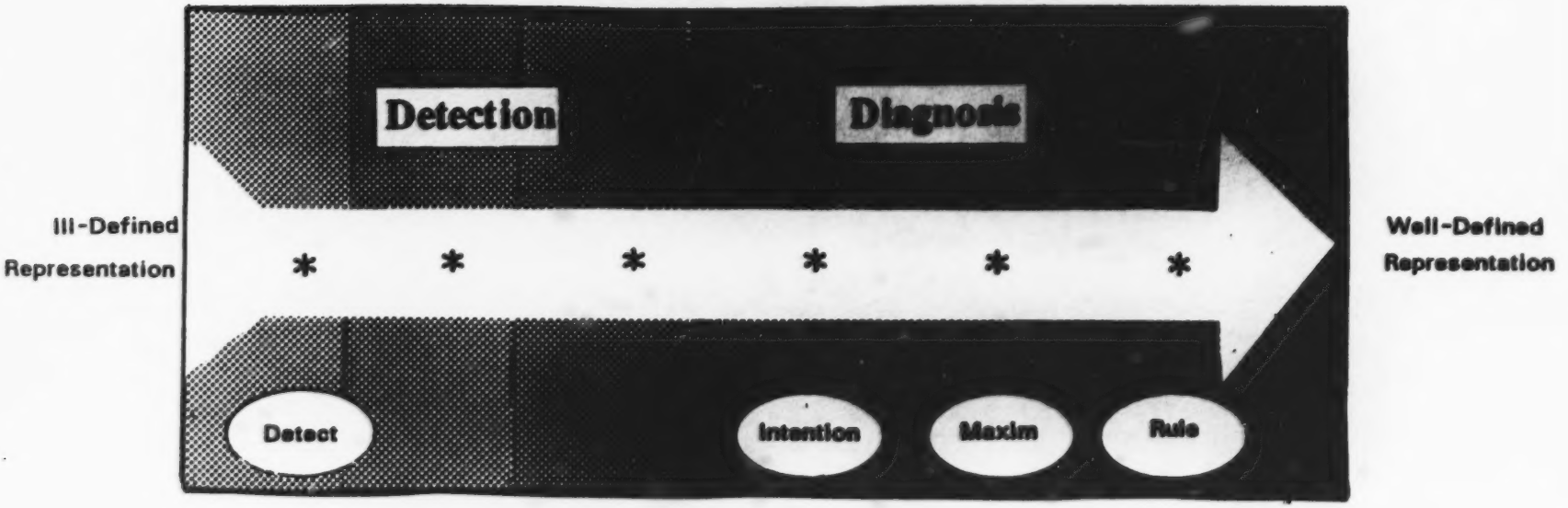


Figure 9. Problem Representation Continuum.

| EXPERTS | NOVICES |
|---|---|
| 20 | 6 |
| 18 | 6 |
| 10 | 3 |
| 9 | 2 |
| 9 | 2 |
| 5 | 1 |
| 2 | 1 |
| Avg. 10.4 | 3 |

Significance of Difference $p < .01$

Figure 10. Number of Global or Whole-Text Problems Diagnosed

| Subjects | Total Diagnoses | Type of Diagnosis | | | |
|---|---|---|---|---|---|
| | | Rule | Maxim | Intentional | Multi-Sentence Intentional |
| EXPERTS Instructors | 324 | 14 | 80 | 230 | 82 |
| % of Total | | 4 3% | 24 7% | 71% | 35 6%* |
| n=4 | | | | | |
| Professional Writers | 107 | 18 | 19 | 70 | 13 |
| % of Total | | 16 8% | 17 7% | 65 4% | 18 6%* |
| n=3 | | | | | |
| TOTAL EXPERTS | 431 | 32 | 99 | 300 | 95 |
| % of Total | | 7 4% | 23% | 69 6% | 31 6%* |
| AVG | 61 6 | | | | |
| TOTAL NOVICES | 214 | 15 | 64 | 135 | 14 |
| % of Total | | 7% | 29 9% | 63 1% | 10 3%* |
| AVG | 30 5 | | | | |
| n=7 | | | | | |

* % of Intentional Diagnoses
Expert/Novice Difference in Total Diagnoses Significant at 05

Figure 11 Summary Data on Total Diagnostic Comments

| | REVISORS | | REWRITERS | |
|---|---|---|---|---|
| | Expert | Novice | Expert | Novice |
| | 160 | 61 | 34 | 18 |
| | 71 | 35 | 31 | |
| | 64 | 33 | 12 | |
| | 59 | 29 | | |
| | | 26 | | |
| | | 12 | | |
| Avg. | 88.5 | 32.7 | 25.7 | 18 |

Figure 12. Revisors vs. Rewriters : Total Number of Diagnostic Comments

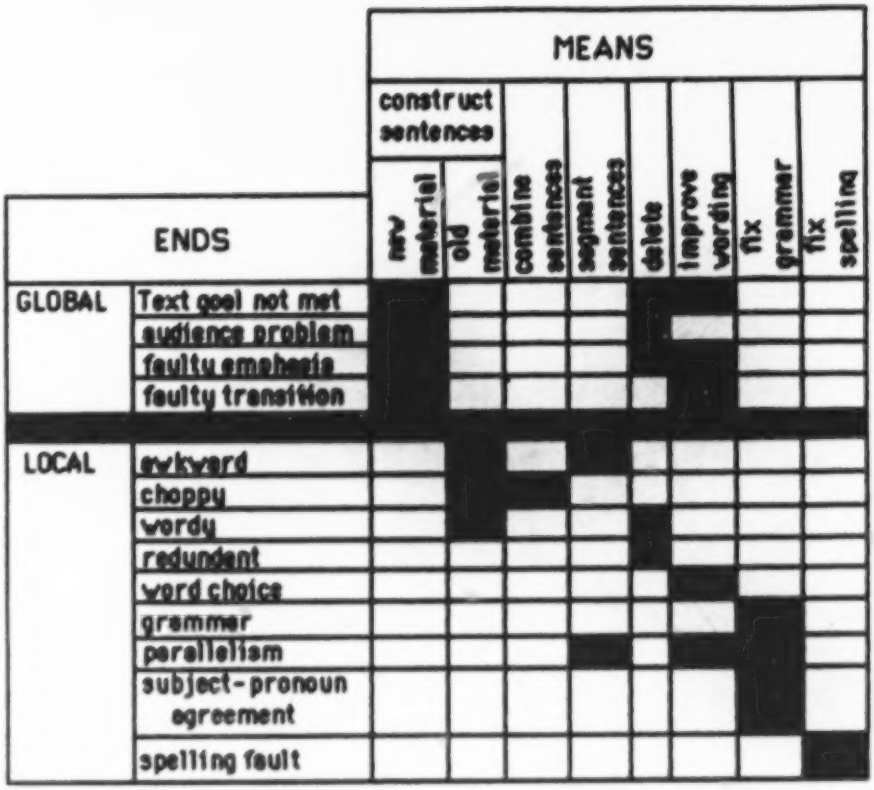


Figure 13. A Sample Means-Ends Table from a Novice Revisor.

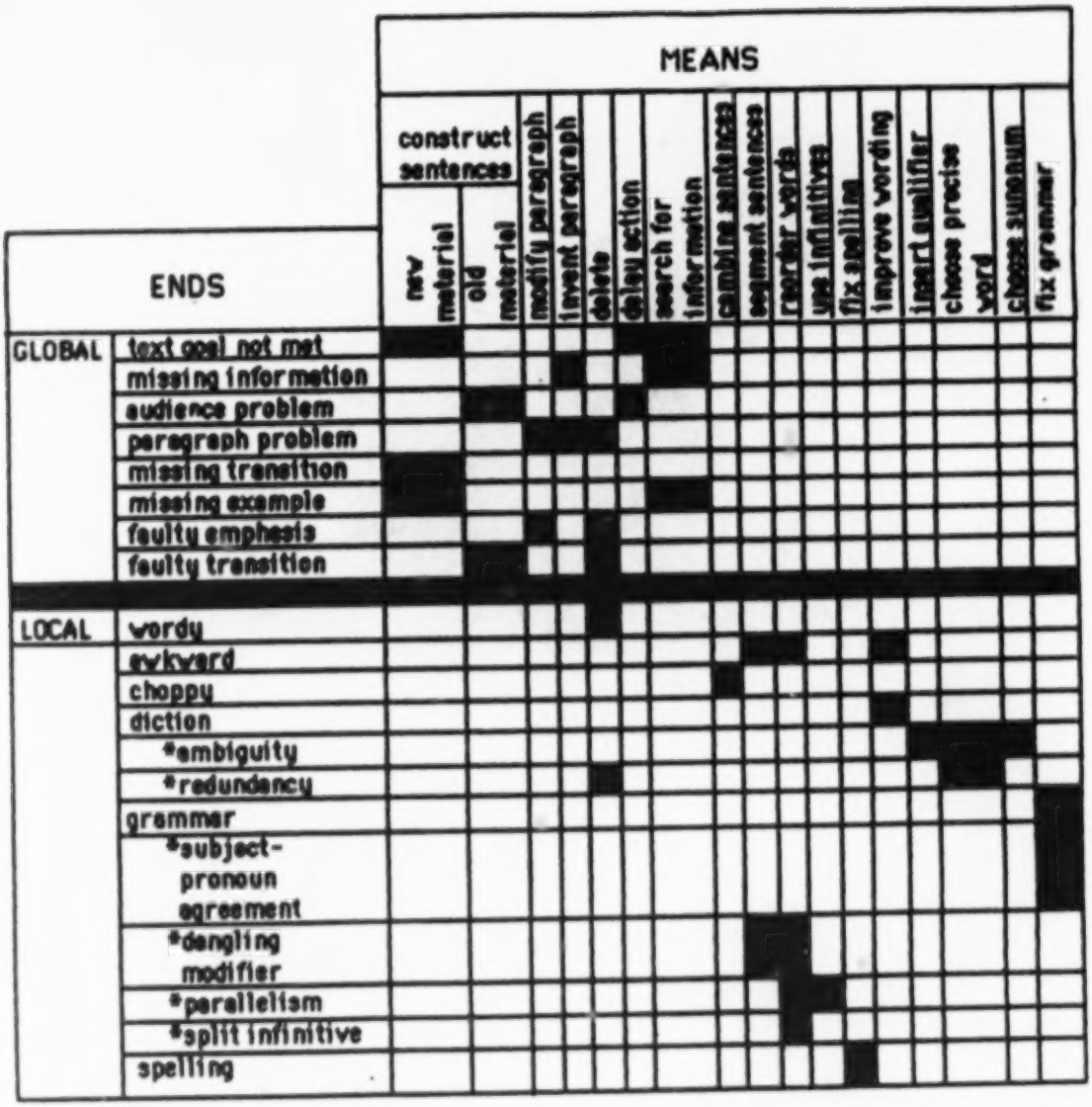


Figure 14. A Sample Means-Ends Table from an Expert Revisor.

APPENDIX A

INSTRUCTIONS FOR REVISION TASK

Here is a copy of a letter written by the Director of Women's Athletics at Illinois Central College, to Jane Fisk, Director of Women's Athletics here at CMU. The letter is in response to a recent article in Focus, CMU's faculty newspaper, on women's athletics at CMU. The author, who is an old friend of Jane Fisk's, congratulates her on her field hockey team's performance, discusses the Focus article, and gives Jane suggestions for writing a brief, one-page handout for freshman women on women's athletics at CMU. Jane decided to take her up on her suggestion. She wants to stick as closely to Janet's discussion (in the marked section of the letter) as possible. However, the PR person for the Physical Education Dept. has forcefully suggested that it needs some work on both the writing and the overall presentation.

Directions:

Your job is to read the letter and revise the marked-off section so that it can stand on its own as a one-page handout. It will be titled "Misconceptions and Facts about Women's Sports at CMU." (the title the author of the letter suggested) and placed in all freshman women's mailboxes.

In this task, you are to stick to the prose just as much as you can, while editing it to improve it. The typist will be given the letter, with your changes marked on it, along with any additional sheets of paper you happen to use. Be sure to mark clearly in the letter where any additional

words, phrases, sentences, etc. belong so the typist knows where to insert them when making a final copy. Also, when you are revising, do not erase any changes you make; instead, whenever you wish to get rid of text, simply draw a line through it.

REMEMBER: YOU ARE REVISING A ONE-PAGE HANDOUT FOR FRESHMAN WOMEN, STICK TO THE PROSE AS MUCH AS POSSIBLE - DON'T REWRITE, JUST REVISE.

## QUESTIONS SUBJECTS FREQUENTLY ASK ABOUT TALK-ALOUD PROTOCOLS OF THE WRITING AND REVISION PROCESS

### WHAT IS A TALK-ALOUD PROTOCOL?

A protocol is a sequential recording of a person's attempt to perform a task. Developed early in this century, protocol analysis is a powerful tool in educational research. In particular, the information captured in "talk-aloud" protocols enables the educational researcher to construct detailed models of human thinking processes, and in some cases to simulate these processes in a computer program. In short, protocols give the researcher a "window" through which to look at otherwise invisible mental processes that occur from moment to moment. We are concerned in this short explanation with talking-aloud protocols of writers revising a text. Listed below are some more questions that subjects of protocol experiments ask.

### HOW IS A TALKING-ALOUD PROTOCOL MADE?

The procedure is really very simple. The researcher will ask you to talk-aloud while you are revising or rewriting a particular document. You are to say out loud what you are thinking. You are not to worry about speaking correctly, stopping in the middle of thoughts or sentences, etc., but you should try to verbalize as *continually* as you can during the entire time you are at work. Pauses in your talk will naturally occur, but try to avoid them. If the researcher feels you are not talking often enough, he/she may prompt you.

### SHOULD I TRY TO EXPLAIN HOW MY WRITING PROCESSES WORK, OR HOW I WOULD USUALLY DO THIS TASK?

Subjects who ask this question are usually trying to do the researcher's work themselves, at the same time they are revising or rewriting. You are not to describe what you "would" do, but *only what you are actually thinking about at the time you are working*. In fact, you are *not* expected to "analyze" your writing habits or creative processes at all. You are not being asked to "introspect," or to give an explanation or interpretation of your writing. You need only say what is on your mind at the moment. Concentrate on the task you have been given, and simply say aloud whatever occurs to you.

**HOW CAN A PROTOCOL CAPTURE MY THINKING PROCESSES IF I CAN'T SAY ALOUD EVERYTHING I AM THINKING?**

Of course, you will not be able to say *everything* you are thinking when you are completing even a simple writing or revision task. A portion of your thinking is lost and falls between the cracks. But almost everything you do say is valuable to the researcher. The amount of information retrieved from the talk-aloud method probably exceeds the amount to be gained by any other research method currently employed for the study of how people write and revise. Moreover, your transcribed protocol is not the only data studied by the researcher; your finished or revised text is also studied, and compared with the talk-aloud transcript. By itself, your finished text tells the researcher very little about the *processes* you used to create the text. However, when your finished text is "matched" with the protocol transcript, the researcher has a much more detailed picture of how your writing and revising unfolded. Again: concentrate on the task and on whatever you are conscious of as you work. Say aloud everything that comes to mind.

**DOESN'T TALKING-ALOUD INTERFERE WITH MY THINKING, SO THAT I AM NOT THINKING AND WORKING AS I NORMALLY WOULD?**

This question is often asked, and rightfully so. It's a very important question for researchers to deal with. At the present time, no one knows for certain if talking-aloud does interfere with your thinking during problem-solving. A lot of research is presently being conducted to find out. So far, researchers have been unable to find any strong evidence that talking-aloud interferes with thinking. Some research has even shown that, with very little practice, you can solve the same problem in the same amount of time whether you are talking aloud or not. Talking aloud can also improve decision-making, and many people talk to themselves when they write anyway. The first few minutes of a protocol may feel awkward, but with a little practice this feeling will disappear and you will feel more comfortable.

**SHOULD I WRITE AND TALK AT THE SAME TIME, OR ONLY BEFORE OR AFTER I WRITE SOMETHING DOWN?**

You should *talk as continuously as possible*, whether you are writing or not. Sometimes you will find yourself only able to say exactly what you are writing on paper. This is perfectly fine, so long as you don't pause too long between words. If you do, the experimenter will prompt you.

**SHOULD I TALK-ALOUD EVEN IF I AM JUST RE-READING WHAT I'VE WRITTEN?**

Yes, you should. Avoid the temptation to mumble if, and when, you re-read your text. Even if you are skimming rapidly, and not re-reading sentences in their entirety, talk-aloud and make sure your voice is audible and clear.

**SAMPLE PROTOCOL**

The experimenter will now play a tape-recording of a person talking aloud as he works. The tape illustrates the concerns raised above. In particular, notice how the writer-subject concentrates on the task: for the most part he talks *while* he both reads and writes, though the experimenter has to prompt him a few times to keep talking. In this first sample segment, the *italicized* parts indicate what the writer-subject is actually putting on paper as he talks. The "Directions" he is reading are set off in single quotation marks:

"um - ok - my task is to - let's see you want me to read the directions - ok - Directions Your task is to describe the drawing you are given so accurately that someone could reproduce it only from your written account. Remember, however, that efficiency is as important as accuracy. - ok - I get it - this isn't easy because it says accurately - how accurately ? - and efficiently - the detail that's tough - ok - I guess I'll get started and make an outline - no, wait a minute - reproduce they say - does that mean it just has to have the same parts, but not look the same - completely the same? - says so accurately that - uh - could reproduce it simply from your written account - (please keep talking, ok?) - yeah, ok - I just don't quite understand what these instructions are trying - it seems there's probably a hidden problem in here - I guess we'll - the first thing is to just list the all the parts - ok, here's a castle, kind of a snake, a star - (you're falling silent; please keep talking) - I keep thinking about the assumptions in the directions and its like - it looks like - I wonder if I could just sneak some pictures into my writing - so the guy, the reader who's doing - uh, no, that would be cheating I guess - how about if I call this shape like an open book, yeah, it does look like a book, or the corner of a room - and then describe each of the walls as pages in turn - yeah, that's a snap - just start off writing for now and go back later to see - *Envision a open book - In this drawing the left portion of the book* - no scratch that - of the *open book looks like a rectangle, but the top and bottom lines of it slant inward, toward the middle of the paper* -uh, just to get this exact, uh - left to right - ok - *left to right* - is this confusing? - I don't think I like talking about the middle of the paper -

**********

Here is more of the same writer-subject's protocol, a little later on. In this segment, notice that the underlined parts of the protocol indicate when the writer-subject is reading or re-reading his text. The *italicized* text here indicates when the writer is revising parts of his text and adding in new parts.

"well now what a job - that does it for paragraph one I think - it's not very good I think but at least it's exact - they have to get the biggest most general things first so they'll know where to put others - this - all this left and right seems hard but - it seems like this needs to be changed here - (you're falling silent again; please keep talking) - I'll read to see how this sounds - it sounds too complicated, anyway, well, let's see - Envision an open book - in this drawing the left portion of the open book looks like a rectangle but the top and bottom lines of it slant inward, toward the middle of the paper from right to left - Ok, ok, that's clear enough, not perfect, but - The right portion looks like a trapezoid - - maybe can't say that - can't say trapezoid because some people may not know what that is exactly and this has to be very simple so a child can do it - ok so like a *square but its top and bottom lines are also slanted inward toward the center of the paper from right to left*- well that's wordy but I don't have to worry about the term now - the slants from the rectangle meet with the slants of the square in the center of the page at a vertical line - no, change that - meet with the slants *from* the square in the center of the page at a vertical line - this vertical line is the rectangle's right side and the trapezoid's - uh, nope, trapezoid again, have to change that to square - so -this vertical line is the rectangle's right side and the - *the square's* left side - there, I think that's a lot better, a lot - so now I'm ready to describe the left inside cover of this book -or so-called book - I think I'll re-read this from the start again -

ILLINOIS CENTRAL COLLEGE

COMMUNITY COLLEGE DISTRICT NO. 514
P.O. BOX 2400 • EAST PEORIA, ILLINOIS 61611 • TELEPHONE: (309) 694-5011

December 2, 1980

Jane Fisk
Director, Women's Athletics
Dept. of Physical Education
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dear Jane,

Congratulations to you and the field hockey team on this season's performance!

I was quite pleased to see that piece in the recent Focus (Vol. 10, No. 3, November 1980) on the situation for women athletes at CMU. It doesn't say everything that probably needs to be said on the subject, but it's a beginning! I think you're right about student pressure being needed to create change. For myself, I have found that the only way to get a large, active group to lobby for more funding is to recruit more women for the teams. That may seem like getting the cart before the horse, but I think it wouldn't be difficult to attract more women if a little more PR was done.

Most women are reluctant enough to participate in sports, let alone busy CMU students. I've found that many women seem to have a lot of misconceptions about the time required, necessary ability, etc. Our department found it very useful to make a brief handout giving the reasons women should get involved -- I think it was titled "Misconceptions & Facts About Women's Sports at Illinois Central College." It was especially effective with the incoming freshmen, and a comparatively inexpensive method of recruitment. We used all the handouts and I seem to have misplaced the original, but the ideas we included were roughly the following:

BEGIN TO REVISE HERE

Many naive women students possess the assumption that it is necessary that th be superlative athletes in order to successfully be a member of a varsity team. T fail to realize that schools like Illinois Central (CMU) are Division 3 AIAW schools--that is, schools that do not offer athletic scholarships. Therefore, many of the teams are less competitive than many students expect. In sports like fencing for a long time many of our varsity team members had no previous experience anyway. Even some students without background if they have good

coordination have learned so much in their first semester that they even surpassed other with a background in athletics in high school.

I don't want to infer that the only chance women get for participating in sports is on varsity teams. Intra-mural sports are not quite the same as varsity sports, in which the rules are better, equipment is better, and with the techniques of the players being more developed.
Irregardless, IM sports may be the choice for many women--they can be just as much fun and take less time.

Perhaps a more influenctial aspect in freshman students' reluctance to participate in sports is one's fear of the time committment. Many new students, worried about the academic demands of college and inexperienced at organizing their time, feel that participating in sports is another pressure they don't need. It is hard for them to see that the benefits of sports, are worth the time invested. However, being involved in sports demands that the student develop their ability to organize their time. Its structure encourages the student to plan their time. And being part of a varsity team, very fine freindships are are a further dividend for the student athlete.

Students also gain many personal benefits from an involvement in sports. They develop cardio-respiratory endurance and freshman athletes avoid that common enduring affliction--freshman-year weight gain. Students also learn to interact with people, cooperation, punctuality, responsibility, how to win and lose correctly, and how to treat your opponents. All these abilities are important in the working world--which is why potential employers are impressed with a candidate who has experience in athletics.

There is, however, one comment which I frequently hear and which is more convincing than all these arguments, and that's a comment I hear time and time again from women who've started sports late in their college career--"Oh, I wish I'd started earlier!"

These arguments tend to make a very convicing little handout: we had about 25% more freshmen sign up for teams the year we used one than had signed up the previous year. By the way, you should remember to list the sports women can participate in at CMU. ] END REVISION HERE

I hope this method of recruitment works as well for you as it has for us. I'll send you a copy of the handout if the original turns up. Good luck, and best wishes for your teams' continued success.

Warmly,

Janet Grotzinger
Director, Women's Athletics